



Digitized by the Internet Archive
in 2012

Motivation and Market Behavior

MOTIVATION AND MARKET BEHAVIOR

Edited by

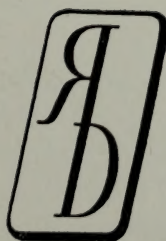
ROBERT FERBER, *Research Professor of Economics*

and

HUGH G. WALES, *Professor of Marketing*

Both of the University of Illinois

Prepared under the Sponsorship of the
AMERICAN MARKETING ASSOCIATION



1958

RICHARD D. IRWIN, INC.

HOMewood, ILLINOIS

© 1958 BY RICHARD D. IRWIN, INC.

ALL RIGHTS RESERVED. THIS BOOK OR ANY PART THEREOF MAY NOT BE
REPRODUCED WITHOUT THE WRITTEN PERMISSION OF THE PUBLISHER

First Printing, January, 1958

Library of Congress Catalogue Card No. 58-5803

PRINTED IN THE UNITED STATES OF AMERICA

FOREWORD

This compilation of readings on MOTIVATION AND MARKET BEHAVIOR has been sponsored by the American Marketing Association for very specific reasons. First, much of motivation research has been directed toward an explanation or prediction of human behavior in regard to the purchasing function. Therefore, the methods employed and the conclusions reached are of great interest to marketing practitioners and teachers. Second, there is considerable controversy at present about what should be included under the term "motivation research," whether it is something new or only a logical extension of research work of the past, whether the methods used to discover motivation are sound, and whether confidence can be placed in the results of investigations in which these methods are used. Apparently confusion prevails in regard to these questions. Comments on motivation research range all the way from a hearty endorsement of methods and results to a considerable measure of skepticism on both. The truth lies somewhere in between these extreme positions and clarification is therefore needed.

The American Marketing Association is dedicated to a scientific approach in the study of marketing problems. Such an approach simply means that painstaking effort is used in determining the relationship between phenomena which affect the movement of goods and services between producers and consumers. The Association is also dedicated to the extension of research in the field of marketing and to the dissemination of information which will serve an educational purpose and ultimately improve marketing methods.

In this particular instance, the objectives of the Association have been well served by bringing together in one volume a carefully selected group of articles on motivation research. Through the publication of such a volume it is believed that readers can more easily secure a sound, well-balanced, critical view of the present status of motivation research and, particularly, in reference to market behavior. This objective, of course, could only be attained if the articles were selected in such a manner that both sides of controversial issues were clearly presented. Those who reviewed the manuscript for the Association prior to its publication believe that the editors have selected the articles wisely and

with the need for objectivity in mind; and that, in so doing, they have served the Association and readers of its sponsored publications well.

Ann Arbor
August 6, 1957

D. MAYNARD PHELPS, *President*
American Marketing Association

PREFACE

The search for motivations in market behavior is, on the one hand, one of the most widely discussed—and exploited—subjects in the field of marketing and, on the other hand, one of the most difficult subjects about which to obtain adequate literature. This is due in part to the interdisciplinary nature of the subject, as a result of which pertinent material is likely to appear in a myriad of sources—ranging from journals on clinical psychology to advertising newspapers—so that only the specialist in the field is likely to come across any appreciable proportion of the total. The difficulty of securing adequate literature is enhanced by the highly controversial nature of the subject, as evidenced by the continual debates regarding the meaning of “motivation research.” On almost any controversial subject it is much more easy to obtain partisan views favoring one particular side or approach than to obtain a range of different interpretations. This is particularly true when so many different interpretations are possible, as is the case with the analysis of motivations.

The object of this volume is to provide a wide range of interpretations for the benefit of marketing and business people. In effect, it aims to present, first, different interpretations and attitudes toward the study of motivation, and second, various approaches to solving problems in human motivation. The basic perspective toward the study of motivation taken in this volume is a very broad one, namely, the search for answers to the “why” of human behavior.

Physical behavior may consist of a simple act, but the motivations underlying it are as often as not highly complex. Even from the individual point of view, seldom is an act brought about by a single motivating factor. Rather it is likely to represent the balance of numerous forces, some distinctive to the individual and others more of an external nature, such as the weather. To identify the “real” motivating factors for many acts is an almost hopeless task because of the existence of innumerable levels of motivation.

The problem becomes enormously complicated when we turn to market behavior, where individual motivations become intermingled with social forces. For this reason, and because of our market orientation, we must be concerned not only with the “whys” of individual behavior but also with the “whys” of aggregate behavior, and hence with

techniques that can be applied to aggregates as well as to individuals. Essentially, from a marketing point of view, the analytical task becomes to identify those motivations pertinent to the problem at hand, and preferably those motivations about which corrective action can be taken.

The search for these motivations involves the use of all sorts of techniques—from statistics, from psychology, from economics, from sociology, and from other social sciences. Some would include all of these techniques under the heading “motivation research,” while others would include under this heading only those psychoanalytical methods aimed at probing one’s emotions and inner feelings. Since the term “motivation research” has come to mean so many different things to different people, it seemed best to employ a different title for this book, one encompassing “motivation research” in its varied meanings, as has been done.

By taking such a broad perspective, it is possible to indicate the many different approaches that are available in dealing with “why” questions as well as the fact that a dimension of motivation that may be useful in one case may not necessarily be of any value in another. Recognition of the existence of different levels of motivation and knowledge of when and how to carry out an analysis on each level is indispensable to a sound understanding of the subject.

The very existence of these numerous levels of motivation, in addition to the economic forces prevailing on any market, highlights the problem of validation. The fact that sales may have increased when a certain promotional approach is used does not of itself prove the effectiveness of that approach unless the influence of other major factors on sales has been first taken into account. Unfortunately, this is rarely done or even attempted in motivation studies. It is, of course, invariably difficult to carry out such “true” validation tests, but the sooner such attempts are made the sooner the barrier will be lifted to major new developments in this area.

The organization of this volume is closely identified with this broad motivation approach. General discussions of the meaning and nature of so-called “motivation research” are provided in Part I. This part brings together some of the most provocative writings of the proponents and opponents of “motivation research” as viewed in its more narrow, psychological sense. It also provides a firm basis for grasping the many different dimensions of motivation that can exist in any given situation.

The other three parts of the book present illustrations, from a wide variety of sources, of different approaches that have been used to explain market behavior. Part II is devoted to the psychological approach, with

projective techniques receiving the main emphasis, for some of the principal controversy in the field concerns their validity. Hence an attempt has been made, as in Part I, to provide views both for and against.

Nonpsychological survey techniques are the subject matter of Part III. The selections in this part attempt to provide an idea of the variety of such techniques that can prove useful in relating motivation to market behavior, ranging in complexity from simple analysis of percentages to highly refined scaling methods.

There is some tendency to overlook the fact that survey techniques are not the only means of explaining why people act as they do. Part IV is designed to illustrate how techniques of aggregate analysis can be, and have been, brought to bear on this problem. Of particular interest should be the great variety of techniques covered in this part of the book.

The bibliography at the end of the book provides the reader with a brief summary of the main references listed in each article contained in this book. In addition, the bibliography provides some additional references on motivation techniques thought particularly useful by the editors.

Although this is a symposium volume, and hence one not likely to be blessed with much continuity, several attempts have been made to remedy this deficiency. For one thing, introductions to each part of the book attempt to place the material in that part in proper perspective in relation to the material in other parts of the book. Second, headnotes at the beginning of each piece discuss its general nature and show how it relates to other material in the book. Third, various changes in style and even at times in means of expression have been carried out to synchronize the writing of the different authors. In all cases, this has been done with the approval of the authors and in no case have the changes been so drastic as to eradicate each writer's distinctive style. Fourth, an attempt has been made to make the format of the various pieces completely uniform, such as in grammar, layout of headings, and presentation of tables and charts.

The editors of this volume would like to express their great appreciation to the many individuals and organizations who gave permission for reproduction of their works. Their identities are given in the initial footnotes to the various articles. Thanks are also due to Elaine Paden for her excellent editorial work in putting the material together, to William C. Gordon, Jr. and the office staff of the American Marketing Association for the fine co-operation provided in helping to prepare this volume for the press, and to Edward McGarry for his critical reading of the manuscript.

In closing, it deserves to be stressed that the selection of the articles contained in this volume, and the interpretations placed on them and on motivation research in general, do not necessarily reflect the views of the American Marketing Association. These decisions were entirely at the discretion of the editors and it is toward them that any criticism should be directed.

ROBERT FERBER
HUGH G. WALES

Urbana, Illinois
August 1957

TABLE OF CONTENTS

PART I—GENERAL APPROACHES

Introduction	3
1. A Psychologist's View	4
<i>Charles Cannell</i>	
2. Advertising Strategy and Theories of Motivation	11
<i>Wroe Alderson</i>	
3. Toward an Understanding of Human Behavior	21
<i>Ernest Dichter</i>	
4. How Psychiatric Methods Can Be Applied to Market Research	31
<i>James M. Vicary</i>	
5. A Case Study: What Automobiles Mean to Americans	36
<i>Pierre D. Martineau</i>	
6. Motivation Research—Opportunity or Dilemma?	50
<i>Alfred Politz</i>	
7. Rationality and Irrationality in Motivation Research	64
<i>L. Edward Scriven</i>	

PART II—PROJECTIVE AND RELATED PSYCHOLOGICAL TECHNIQUES

Introduction	75
8. Principal Psychological Techniques in Motivation Research	76
<i>Lyndon O. Brown</i>	
9. Some Hazards of Motivation Techniques	89
<i>Steuart Henderson Britt</i>	
10. An Application of Projective Techniques	93
<i>Mason Haire</i>	
11. Projective Techniques in Attitude Measurement	103
<i>Irving R. Weschler and Raymond E. Bernberg</i>	
12. Can the Clinical Techniques Be Validated?	122
<i>Darrell B. Lucas</i>	
13. Projective Techniques from an Analytical Point of View	133
<i>Robert Ferber</i>	
14. The Thematic Apperception Test	143
<i>Gardner Lindzey</i>	

15. The Error-Choice Technique	170
<i>Kenneth R. Hammond</i>	
16. The Rorschach Test	182
<i>Daniel R. Miller</i>	
17. Cartoon and Picture Devices	197
<i>Martin Zober</i>	
18. The Meaning of Gasoline Symbols	206
<i>William Henry</i>	
19. Play Techniques for Interviewing on Durable Goods	231
<i>C. W. Garber, Jr.</i>	
20. A Sociopsychological Approach to Consumer Behavior	233
<i>Warren J. Bilkey</i>	

PART III—NONPROJECTIVE SURVEY TECHNIQUES

Introduction	247
21. A Qualitative Approach	248
<i>Allan R. Wilson</i>	
22. Measuring Consumer Attitudes toward Products	268
<i>George H. Brown</i>	
23. Why People Buy Particular Brands	277
<i>Seymour Banks</i>	
24. Measuring the True State of Opinion	293
<i>Louis Guttman</i>	
25. A Statistical Approach to Motivation Problems	315
<i>J. B. Lansing</i>	
26. Why Television Commercials Succeed	321
<i>Horace S. Schwerin</i>	

PART IV—OTHER TECHNIQUES

Introduction	337
27. Why Consumer Purchases Change over Time	338
<i>Ruth P. Mack</i>	
28. Why Advertising Readership Varies	348
<i>Dik Warren Twedt</i>	
29. What Makes a Best Seller?	361
<i>John Harvey</i>	
30. Why Daily Department Store Sales Fluctuate	381
<i>A. T. Steele</i>	
31. Why a Mail-Order Catalogue Is Successful	388
<i>George Kingsley Zipf</i>	
Bibliography	403
Index	431

PART I

General Approaches

INTRODUCTION

The first part of this volume is devoted to selections on the relative merits of different approaches to the study of motivation. The discussion is centered largely on the validity of the psychological approach to the subject, the matter about which much of the controversy is concerned. Most of the discussion is in general terms; the more technical material on specific methods has been relegated to later parts of the book.

To supply proper perspective on the subject, an attempt has been made to include not only extreme viewpoints but also samples of the more moderate approach. The first selection, by Charles Cannell, is one of the latter type, presenting a thought-provoking discussion of the contributions that the psychological approach can make to marketing analysis. Following it is an equally interesting article, by Wroe Alderson, and with yet a more concrete approach, discussing the implications of different psychological theories of motivation to advertising strategy.

More positive views espousing the psychological approach are presented by commercial psychologists Ernest Dichter and James M. Vicary, who are currently perhaps the strongest advocates of this type of approach to the study of motivation in marketing. Hence, these two articles represent in many ways the extremes of this point of view.

The following article, by Pierre Martineau, is for the benefit of the reader who may feel at this point that he has had enough reading about theoretical advantages and would like to see how these techniques are applied in practice. This study, illustrating how projective techniques were employed in ascertaining the factors underlying people's selection of particular automobiles, should serve well in this regard.

Finally, it is only fair to present some critical views of the psychological approach to market motivation studies, and no better representatives of this viewpoint would seem to be available than writings from Alfred Politz Research and A. J. Wood and Company. Both of these organizations have vociferously opposed the psychological approach in clear, simple terms. The selections reproduced in Part I are felt to reflect about the best-written and most persuasive of these discussions.

1. A PSYCHOLOGIST'S VIEW*

This article helps place in proper perspective the role of psychology in marketing. It provides a clear and simple explanation of the different psychological approaches to the study of motivation, and relates these approaches to the study of market behavior.

Particularly noteworthy is the emphasis on the study of perception as a necessary complement to motivation and on recognizing the existence of different levels of motivation in any given market situation. In effect, the latter is an elaboration of the importance of validity tests in conducting motivation research.

People seem always to have been interested in problems of motivation. They have tried to predict what other people will do, sometimes merely from curiosity, and sometimes because, for one reason or another, people have wanted to explain and control the behavior of others. From early times, when the explanation of motives was that people were either possessed of devils or were merely following a preordained plan, man has been interested in understanding the actions of others.

Researchers in professional fields such as criminology, social work, psychiatry, psychology, and many others are trying to understand human motives, to predict how people will act, and to figure out why they act as they do. We see now a great upsurge of interest in motivation on the part of those engaged in marketing. It may be erroneous to say that this is a new interest. Perhaps marketing simply has discovered a new term, "motivational research," for what has been pertinent all along.

Motivation has always been a major field of study for psychologists. There are many methods of investigation and many theories to explain human behavior. These range from instincts, or motives which are inherent in the individual, to studies of neural mechanisms, to social or environmental conditions as the moving forces. To see how the modern psychologist looks at motivation, let us review some of the major points of view in psychology and then consider how some of these principles may apply to the marketing field.

* Adapted from a talk, "A Psychologist Looks at Marketing," by Charles Cannell, Survey Research Center, University of Michigan, given at a meeting of the American Marketing Association, University of Miami, April, 1955.

PSYCHOLOGICAL APPROACHES TO MOTIVATION

The Laboratory Psychologist. Take, first, the major field of the laboratory psychologist. The laboratory psychologist generally uses as experimental subjects white rats and other animals rather than humans. He does this because the motivational pattern of the animal is less complicated and therefore more easily accessible to experimentation and measurement. This approach is the oldest in psychology from the point of view of motivational research and has provided much research data and many theories which are applicable to human motivation. In general, the experimental psychologist's approach to motivation has tended to focus on the biological drives, or physiological tensions, or "body needs." He emphasizes that much of human behavior is motivated primarily by biological drives, such as hunger, thirst, and sex. A typical early experiment with white rats uses a Y-maze. The rat has learned that at the end of one branch of the Y is food, and at the other branch of the Y is water. By depriving the rat of both water and food, one can determine which of these drives is stronger at any particular time, simply by observing which pathway the rat chooses as he runs the maze. Similar experiments have been repeated many times using various drives to see which are strongest.

It is sufficient here merely to recognize the importance of the internal biological nature of motivation which the laboratory psychologists have emphasized. It appears that this is the basis of all motives, that as society becomes more complex biological drives also become complicated, being changed and influenced by social processes. But basically most motivation is due to the physiological tensions set up in the organism which cause the animal, be it rat or human, to behave in a particular way.

The Clinical Psychologist. A second major area in psychology which has been deeply involved in the study of human motivation is clinical psychology. The clinician emphasizes the influence of psychological factors as opposed to the biological factors just mentioned. It is not that he regards the biological factors as unimportant in determining motivation, but because at the psychological level he can best control, influence, and change motivational patterns. The field of clinical psychology, of course, has been strongly influenced by psychoanalytic theory. Psychoanalysts have emphasized the problems of the child learning to handle internal biological drives, particularly the sex drive, within the norms of our society. They point out the

conflict between motives and the socially acceptable behavior. Thus, for example, a child can experiment with the hunger drive by eating unusual foods or even eating sand, and his behavior is more or less condoned. If he experiments with sex, however, the cultural norms rapidly catch up with him, and such experimentation is punished.

Conflicts between basic drives and the norms of society give rise to severe stresses on the individual. In highly oversimplified terms we can describe what happens as a result of such stress. Since the conflict cannot be tolerated indefinitely, he "represses" or "forgets" the conflict by forcefully driving it out of consciousness. The conflict then remains in the unconscious. Even though the motive itself is now unconscious it is still a powerful force, but one of which the person is unaware. Repressed or unconscious motives show up in peculiar ways. Thus, the clinician is aware of the fact that the reasons a patient gives for a particular action are very likely not the real or basic motives for his particular behavior. The clinician has developed various techniques, some of which will be discussed later, to attempt to uncover and study the motives which are below the surface of consciousness and which the patient is unable to report, or is unable to report accurately. To take an example, one case in the literature tells of an aviation student who was basically a proficient flyer, but each time his flying instructor gave him particular instructions the young man appeared unable to comply and usually did the opposite. This puzzled both the instructor and the young man until, in the clinical interview, the student recognized that he was reacting to the instructor in the way he had reacted to his domineering father. Even though he wanted very much to be an aviator, the unconscious motives leading to his reaction to the instructor were such that they hampered his ability to fly.

To summarize briefly the major influences of clinical psychology in motivation: (1) An individual is often quite unaware of his own motives; (2) motives are complex and very often conflicting, a single act seldom being the result of a single motive; (3) the basic personality structure of the individual is important as a seat of motives, and (4) motivations are found at various levels of consciousness; some can be reported accurately, but many cannot.

The Sociopsychologist. The third major area of the psychological study of motivation can be termed the "sociopsychological" approach. This is best exemplified in the work of the late Kurt Lewin. The general approach, as the name implies, emphasizes the fact that man is always reacting in a social environment, and that his motives can

be understood only as we understand the environment in which he is reacting, as well as his motives. For example, if a child is placed in a room with a box of candy, it is easy to predict that the child will take some of the candy, unless he happens to have already helped himself several times and is thereby satiated. The child is motivated to eat the candy, and there is a simple motivational pattern. Suppose, however, that the mother just before she left the room warned the child that he should not eat any of the candy. There is now a conflict in motives. The candy is a desirable goal for the child; he is motivated to get the candy. He has, however, the strict warning of his mother not to eat the candy. Can we now predict how the child will act? The answer will depend upon the strength of his desire for candy and the strength of the warning which his mother has given him. Unless both of these are known, it will be impossible to predict just what the child will do.

In this example it can be seen that there is another aspect of the psychological makeup of the individual which we must discern if we are to understand how an individual will react. That other element is the perception of the end results of the act. In the example just cited, we had to know something about what the candy represents for the boy, about the strength of his motive, and the comparative strength of the mother's warning. The sociopsychological approach emphasizes the importance of *perception* as well as motivation.

At the present time the field of marketing seems so impressed with the term "motivational research" that it has sometimes failed to take into account other important psychological factors in understanding behavior. The field of perception warrants the attention of the market researcher. In fact, almost as good a case could be made out for a field of "perceptual research" as for motivational research. Actually, any research or any understanding of behavior must include both perception and motivation.

As another example, a certain man is relatively insecure in his job. He feels that his job is of low status, that he is worth more money than he is getting, and that he is able to hold down a much bigger job than he has at the moment. He feels the need for prestige and he tries by various means to build up his own self-esteem. A short time ago this man was in the market for a car. He finally narrowed his choice down to two. One was a new Ford, and second was a two-year-old Buick. The two cars cost about the same. Can we predict which one of these two cars the man will buy? Talking with him makes it clear that he regards the car not merely as trans-

portation but also as something which will help him to achieve the status he needs. A car is a symbol of prestige to him. It is apparent that his perception of the car is related to his need for self-esteem. Knowing this, we then know more about his motivational pattern and probably can predict that he will purchase the second-hand Buick, which he perceives as a prestige car, rather than the new Ford.

To summarize this sociopsychological approach to motivation, we find an emphasis on understanding human behavior by understanding something about the individual's perception and something about his internal motivation. Knowing these two factors and the conflicting motives with which he is faced, we are better able to understand his behavior.

THE PSYCHOLOGY OF MOTIVATION

Psychology begins with the precept that *all* behavior occurs for a reason, that is, all behavior is the result of motivation. A single act usually involves many motives, some motivating the person in one direction, and some working to motivate the individual in the opposite direction. The study of motivation is, therefore, complicated; really to understand behavior, we must become involved in the complexity of the total motivational structure of the individual. Further, in order to understand why a person acts as he does, simply knowing his motives is inadequate. We need also to understand his perceptions of the situation—the meaning and significance it has for him—so that we may know the basis for his motives.

Motives are both biological and psychological. Motives are internal to the individual in terms of basic drives such as hunger, thirst, and sex, but are drastically influenced and changed by social and psychological forces.

One other aspect of motivation is that the individual is often quite unaware of his motives. In fact, the motivational patterns of the person, if they are accessible at all, often come through in disguised form due to psychological pressures from within the person. This means that often an individual is unable to report his own motives, even if he is willing to do so.

Although many different techniques and methodologies have been employed by psychologists to study motivation, for the sake of this discussion they can be simplified into two general approaches. The first is the clinical case study method.

The clinical psychologist is engaged primarily in analyzing the

psychological life of his patient as a basis for helping the patient to gain insight into his motives. For this purpose, the clinician makes use of depth interviews, projective devices, personality tests, and so forth. It should be noted that these techniques are used by psychologists primarily to learn something about the *individual*, not a group of individuals. In other words, this approach is used most often by the clinical psychologist as a means of diagnosing the problems of a patient. Modifications of these techniques have been used by social scientists to study personality dynamics and personality structure on a statistical basis, but by and large the techniques are employed in the clinic with individual cases.

The second type of methodology used to study motivation is the statistical approach. Here the researcher studies motivation indirectly by testing hypotheses, or by a statistical inference of motives. As an example, during the war studies were made of people's attitudes and behavior toward War Bonds. The research was used by the Treasury Department to develop a more effective program for selling war bonds to the public. One hypothesis was that one of the best ways to motivate people to buy bonds was to ask them personally to buy bonds, that is, that personal solicitation might be much more effective than mass appeals by radio, newspaper, and other mass media.

To test this hypothesis, two cities were studied. In one city, the public was subjected to the usual mass appeals, while in the second city the same appeals were supplemented by an active campaign in which people were asked personally to purchase bonds. At the end of this campaign the bond sales for the two cities were compared, and it was found that the city having made heavy use of personal solicitation had sold many more bonds than the city using only the mass appeal. People in the two cities were interviewed to determine whether they had bought bonds during the campaign and if so why they bought them. The responses in both cities were in terms of helping the soldiers, adding to personal savings, and so forth. Almost no one answered that they bought a bond because someone came to the door and asked them to do so. This is an example of the testing of an hypothesis by an indirect approach. We can infer statistically the type of appeal which resulted from personal solicitation.

It should be clear that in making statistical studies various psychological techniques (such as depth interviews, modified projective techniques, and similar devices) are also available to the researcher. Probably people in marketing research will get more accurate data

from studies of the second type, the statistical approach. Since these techniques can be used with cross-section sampling, the researcher can get a more accurate picture of the public than through the more insightful but limited techniques of clinical studies.

MOTIVATION IN MARKETING

The first job of the researcher in marketing must be to decide at what level of motivation he needs to conduct his research. For example, to know that an individual is motivated to buy a product because of some repressed incestuous desire may be of little use in selling that product. If, for example, we have a problem of whether milk will sell better in square cardboard cartons or in round glass bottles, it may well be that any possible sexual connotations of the bottles are less important than the fact that a square carton fits more easily into the refrigerator. In other words, while deep motivational studies are useful in marketing, it is unwise to assume that in all cases such a study is required.

To look at another example, if a person were asked why he bought a Cadillac car he would likely give answers such as, "It is dependable," "It has a better trade-in value than other cars," or "I like the way it drives." It is quite unlikely that he will give answers like, "It satisfies my need for prestige or status," or "It makes me feel superior to other people." Here is a case where the simple direct question will probably produce misleading information as to the real motives. Now take, as a third example, research on why people do and do not use airplane travel. It may be that the determination of airplane travel has something to do with basic personality characteristics such as personal feelings of security or insecurity. The answers the individual gives as to why he does or does not fly may be less meaningful in predicting which people will fly than in discovering that people who are basically secure will fly, while those who are insecure will not. It must be emphasized that there are various levels at which motivation can be studied. There is no right or wrong level; the decision as to which level to study must be based on the particular task at hand, and the methods adapted to the problem rather than the problems to the methods.

Because the term "motivational research" is such a nice sounding phrase, people in marketing have developed a tremendous interest in research in human motivation. In becoming interested in such research, however, there seems to be a strong tendency to accept *techniques* which the psychologist uses to study motivation (such as depth interviews, projective techniques, and free association tests)

without accepting the psychological *theory* basic to these techniques. Techniques are measurement devices only; they are not theory. Theory is essential to understanding the consumer, and techniques in themselves are useless without theory and without hypotheses derived from theory. People in marketing can gain considerable help from psychology in methods of research, but they should also take advantage of theory and hypotheses relating to motivation and perception in designing and analyzing their studies.

2. ADVERTISING STRATEGY AND THEORIES OF MOTIVATION*

The implications of different psychological theories of motivation on advertising policy are reviewed neatly and succinctly in this interesting article which also contains one of the clearest summaries of the two opposing theories of motivation yet seen by the editors. Such a summary is well kept in mind in reading further selections in this volume. As the reader will readily perceive, most of the psychological work on buyer motivation done during the past few years derives from the psychoanalytic point of view, though Gestalt psychology has recently begun to receive increasing attention.

Motivation research undertakes to meet some of the needs of advertising strategy. The strategist in advertising, as in any other contest, is trying to outguess his opponent by adopting a course of action with a greater payoff. Billions of dollars are spent every year to influence consumer preference for products. There is ample evidence of variation in the effectiveness of the advertising dollar depending on how it is spent. Advertising has come to be the major marketing expenditure for many companies, so that the choice of an advertising strategy is a primary concern for marketing management. In a large company, even a small improvement in the efficiency of advertising may be worth millions of dollars. In a smaller company, it may not seem possible to advertise at all unless some strategic advantage can be found which will give greater effect to each dollar expended.

In trying to devise winning strategies, advertisers necessarily rely upon some theory or explanation of how consumers can be expected to react to products or to advertising appeals. Many advertising experts

* Adapted from an article by Wroe Alderson, partner, Alderson and Sessions, "Advertising Strategy and Theories of Motivation," *Cost and Profit Outlook*, December, 1956.

are inveterate theorists themselves. Often the theories propounded are created on the spot to persuade clients to accept one campaign proposal or another. The great sums at stake and the growing sophistication of both clients and agency executives have created a demand for a more general theory of motivation with foundations in psychology and the other social sciences. Despite real progress in motivation theory and research, the diversity of theoretical positions, particularly in psychology, has created a confusion of counsel. Motivation research is an essential aid to advertising strategy, but the advertising strategist would be well advised to assay the long-range consequences of some of the proposed theoretical positions.

For more than a generation the psychological foundations of advertising theory were relatively simple, consistent, and widely accepted. The behaviorism of John B. Watson was distinctly an American product which seemed well adapted to the American scene and to advertising in particular. According to this view, a child entered the world with little except the capacity to receive impressions and to develop attitudes and habits implanted by its elders. Similarly, the consumer entering the market was like a clean slate on which the advertiser could leave whatever impressions he pleased. This view concerning the role of advertising spread rapidly after Watson himself left the university and entered the advertising field. Endless and massive repetition was regarded as the foundation for advertising success. Habits of buying particular products or brands were to be inculcated in millions of consumers, and constant repetition of simple and forthright messages became standard practice in advertising as in the classroom.

However inadequate this view of consumer psychology may seem today, it has a solid core of truth and continues to be manifested in advertising practice. In its extreme form, this doctrine makes the consuming public an inert and docile mass without the power of rational decision and subject to manipulation at the will of the advertiser. At the same time, it must be admitted that our daily lives are largely made up of useful habits which help us to avoid making an overwhelming number of decisions from moment to moment. A useful habit is not so much irrational as nonrational. It is not opposed to reason but can operate without the active intervention of reason. A rational being can properly make a decision to cultivate useful habits as a way of economizing psychic energy. There are some routines which have the force of habit but never become entirely automatic. This is certainly true of so-called "buying habits." Reminders through

advertising, even though repetitious, can be of service to consumers in maintaining buying routines without really infringing on the prerogative of rational choice.

Motivation research today attempts to penetrate more deeply and to lay the foundation for strategies other than that of conditioning the consumer through massive repetition. This is an inevitable response to the pressures for advertising efficiency. Contemporary motivation research has drawn its inspiration from schools of psychology first developed in Europe as compared to the earlier commitment to the native American school of behaviorism. In fact, nearly all that now goes under the name of motivation research is derived from two great schools of European psychology, each with numerous variations and each now firmly established in the United States. These two schools are in such glaring contrast to each other that the only precept they really have in common is their opposition to behaviorism and its faith in the conditioned reflex. This article will attempt to say something about what these two schools are, about the implications for advertising strategy of adopting one view or the other, and about possible reconciliations between the two for motivation research, for advertising strategy, and for management policy.

CONTRASTING VIEWS OF MOTIVATION

The two principal schools of motivational theory are derived from Gestalt psychology and psychoanalysis. Gestalt psychology is associated with such names as Wertheimer, Kohler, Koffka, and Lewin, all of whom came to the United States in their prime and published some of their most important work here. In fact, Gestalt scarcely became a psychology of motivation until it entered its American phase, having begun as a new approach to the psychological analysis of perception. Psychoanalysis is associated with such names as Freud, Adler, Jung, Rank, Sullivan and Fromm. Psychoanalysis has also gone through a major transition from the preponderantly biological interest of Freud and his followers to the social and cultural viewpoint of more recent writers such as Fromm, Sullivan, Kardiner, and Horney. Gestalt as compared to behaviorism represented renewed interest in conscious mind and rational decision. Psychoanalysis invented and popularized such concepts as the unconscious or subconscious mind. Gestalt as a psychology of motivation is pre-eminently concerned with goal-directed behavior and rational use of the resources of the environment to attain conscious ends. Psychoanalysis, at least in its earlier versions, held that

behavior is primarily determined by instinctive drives and contended that we are unconsciously motivated to seek goals which we do not recognize or may be unwilling to acknowledge even to ourselves.

In a general way, one may be said to emphasize rational behavior and the other irrational behavior, even though it is not always possible to draw a sharp line between these two categories. One definition of rational behavior would be the conscious and deliberate pursuit of goals that are consistent with the survival and well-being of the individual. Psychoanalysis would say that much of human behavior lies outside the area of rationality so defined, and that some of the most fundamental aspects of motivation are hidden below the level of consciousness. To the extent that this is true, it obviously complicates the problem of finding out what people really want or what motivates their behavior. It is also true, however, that rationality of goals or behavior would not necessarily mean that the task of motivation research would be easy. While psychoanalysis holds that ideas are repressed because the ego cannot accept them, people also forget because they have achieved a satisfactory adjustment and have had no reason to recall their original motivations. In case after case there seem to be perfectly practical and common-sense reasons why consumers should prefer one dishcloth, detergent, or depilatory to another. If respondents seem vague when first challenged to explain, it is probably because they have had other things to think about, rather than because they are subject to any great inner tensions or anxieties connected with these everyday products.

One of the difficulties about the concept of rationality is that it is not always considered from the viewpoint of the subject whose behavior is under critical scrutiny. The inherent standards of rationality in a field of consumer use may be quite different from imputed standards of rationality existing in the mind of the outsider. All too often some producer assumes that users place a high value on certain technical characteristics of his product, only to find upon investigation that they have an entirely different conception of its principal virtues. It seems obvious after the fact, for example, that housewives might consider absorbency a valuable characteristic in a dishcloth. Nevertheless, the first draft of one survey questionnaire omitted this topic entirely, even though the manufacturer had made an exhaustive attempt to deal with every characteristic which could possibly interest the consumer.

The two leading schools of thought point to quite different conclusions about the development of personality. Gestalt in its original version pictured the rational mind as endowed with insights which

enabled it to see a solution almost as soon as the problem situation was presented. Later versions make more allowance for learning from experience. The mature personality is one which has become progressively more skilled in the management of the resources of its environment. The mature personality, for psychoanalysis, is one which has finally achieved a degree of poise and balance after surviving nearly disastrous incidents along the way. Growing up, for psychoanalysts of biological bent, is the painful process of recovering from such traumatic experiences as birth, weaning, toilet training, and puberty. Even those with social and cultural leanings picture the typical life history as a continuous battle to master the forces of a hostile environment. Obviously these two views have very different implications as to the way consumers will react toward goods or toward the various appeals presented in advertising. If the first view is correct, the consumer might be expected to regard a product as primarily an instrument for obtaining a given end and to judge it in terms of its instrumental efficiency. If the second view is correct, the consumer might be expected to be much more preoccupied with the symbolic aspect of goods, to utilize them as means of giving vent to suppressed desires, and to be more interested in symbols of mastery than in working tools.

CONSEQUENCES OF THE INSTRUMENTAL VIEW

If goods are working tools or instruments for gaining specific ends, advertising might be expected to take on an educational character. In a service magazine for housewives, for example, the tone of an advertisement might not be too different from that of an article describing a method for dealing with some household product. Like any other teacher, the advertiser might engage in repetition partly to make sure that each subject had learned the lesson and partly because there is a constant stream of new subjects who have not yet been exposed to the lesson. Advertising from this viewpoint sees consumers in their social roles as members of households, as income earners, and as purchasing agents for the household. It talks to them in terms which they can use in talking to their family and friends. Its appeals are made through public channels and are directed to what may be called the "public" life of the individual, in contrast to what is peculiarly personal and private.

To proceed as if consumer behavior is fundamentally rational is to assume that behavior patterns will be convergent. That is to say, consumers with precisely the same problems will tend to adopt precisely

the same solutions after some experience with the various alternatives. Rationality in the form in which it is generally available consists in being able to make comparative judgments among ways of accomplishing the same thing. Thus, at the level of consumer buying, it is reflected in a capacity to learn even when there is not enough creative insight to be right the first time.

Some advertisers who have themselves preferred rational appeals have sometimes been discouraged with the apparent results. Some years ago a leading dentifrice advertiser rebelled against what he considered the fantastic claims made by himself and his competitors. He decided to advertise that the sole function of a toothpaste was to clean teeth and that his product could do it as well as any other. About this time his brand began to suffer a serious decline in sales and he therefore soon returned to less factual and more colorful claims. Interestingly enough, the same manufacturer is today once more treating the consumer as a rational being and is presumably doing all right. Undoubtedly, consumers are steadily becoming more sophisticated, but it is also possible that the manufacturer misread the evidence as to consumer reactions on the first occasion.

It has already been pointed out that the supplier sometimes does not really know what the consumer's problem is. Even if he is familiar with the problem, he may not be clear as to the product features which are important to the user. There is also the possibility that, although there is a patent consumer need, there may not be a widespread conscious recognition of the need among consumers. Sometimes the advertiser must establish the fact that a problem exists before trying to show that his product offers a solution. Emotionally colored language directed to this end is not inconsistent with the postulate of consumer rationality. Absorbed as we may be in our daily affairs, each of us may need to be startled into recognizing the urgency of some problem, the importance of being prepared for the eventuality which it represents, or the frequency with which such occasions may be expected to arise.

Long experience in consumer research gives one considerable respect for the job done by the consumer buyer. She must consider many things bearing on the well-being and happiness of herself and her family and in most instances must choose among a wide variety of products. Some consumers are more effective buyers than others or may have special buying skills in certain fields. These leaders of consumption tend to influence the buying decisions of other consumers. They are likely to talk in practical terms when they recommend a product

to other consumers. Advertising that stresses rational reasons for purchase is more likely to be multiplied through word-of-mouth advertising. The consumer may have mixed motivations for purchase, particularly in the case of an item for her own personal use. Even here she may need rational reasons for justifying the purchase to other members of the household. Advertising that recognizes the postulate of consumer rationality would, therefore, seem to be a basic ingredient for most marketing programs. Such advertising recognizes the social role of the consumer buyer, and is likely to create enduring values for the advertiser since it is consistent with basic trends in consumer attitudes and behavior.

CONSEQUENCES OF THE SYMBOLIC VIEW

Advertising that assigns a large place to the symbolic aspect of goods is related more to the remedial work of the clinic than to the educational task of the classroom. There may be some instances in which the advertiser can afford to play the role of the psychiatrist, but this is a difficult function to discharge through the media of mass communication. The psychiatric needs of the individual are highly personal and the effective symbols may be largely private. While Freud and others insisted that a study of dreams revealed universal symbols, it is common experience that the same symbol can mean many things to many people.

The symbolic approach tends to be divergent just as the instrumental approach is convergent. The very fact that a symbol might be recognized by others may make it less acceptable to the subject as an outlet for secret yearnings. When a product is regarded as an instrument, each user wants to get results as good as his fellows. When a product is regarded as a symbol, he may be more interested in characteristics which will help to set him apart from his fellows. A cherished symbol may either facilitate or delay maturation. It may help the individual to fit an aspect of experience into his life pattern. On the other hand, it may be a means of escape into the realm of fantasy. Thus the advertiser who is especially concerned with symbolic meanings is operating in a field of industrial poetry in which the impact on the audience or the future consequences for the advertiser are hard to predict.

Psychoanalysis divides people into types, but the typologies are almost as numerous as the analysts. Jung talks of extroverts and introverts; Rank, of the neurotic and creative man; Horney, of the compliant, aggressive and detached types. Fromm lists five orientations of person-

ality including the marketing orientation and the productive orientation. There is still the orthodox classification which describes such character types as the oral, the anal, and the genital. One great difficulty in applying such typologies is that no related classifications are available for families or households which, after all, are the fundamental purchasing units. In fact, it seems more reasonable that marketing should start from a study of the organization of household units and then deal with the personality traits which might make for good or poor household organization.

It would be hard to assess the outcome of advertising competition if the emphasis were to be primarily on symbolism. In one instance, a motivation study following the Freudian point of view indicated that consumers preferred vegetable shortening because animal fats were somehow related to a sense of sin. While this might be a consideration favoring vegetable shortening as such, it is hard to see how this point could be made effective for one brand rather than another. Similarly, if it be assumed that there is some symbolic value in lipsticks or cigarettes, all brands would presumably be equally potent symbols.

One justification for this type of approach might be the attempt to achieve variety in advertising copy and presentation. In a given case, anything that can be said as to the instrumental value of a product may have been said many times over. The advertiser may suspect that his potential audience has become bored and inattentive so that these instrumental messages are no longer registering. He might use the Freudian approach as a way of developing new copy slants even though the grosser forms of Freudian symbolism were not actually apparent in the copy. His purpose might be to gain attention and to get new prospects to try his product, hoping that it could demonstrate its instrumental virtues upon trial. Over-all advertising strategy in such a situation would presumably be that of continuing the repetitive messages but making them more palatable through a fresh approach. This type of advertising strategy would really rest upon Watsonian behaviorism even though it gave a passing nod to psychoanalysis.

RECONCILIATION IN MOTIVATION RESEARCH

I have attempted to describe the consequences of these two points of view, each on its own merits. My own preference is for Gestalt as the framework of motivational theory. More than one contemporary psychologist has demonstrated the capacity of the Gestalt position to utilize some of the basic insights of other schools without abandoning

the postulate of essential rationality. One of the most successful is Carl Rogers, who has opened up new vistas in clinical practice through what he calls "client-centered therapy." The final chapter in his book by that name¹ presents, in slightly over fifty pages, what is possibly the most useful synthesis to date of psychological theory for marketing and advertising.

This chapter contains nineteen propositions about human behavior and the processes of adjustment to the environment. The postulate of rationality is embodied in the fifth proposition which reads as follows: "Behavior is basically the goal-directed attempt of the organism to satisfy its needs as experienced in the field as perceived." The organism strives constantly to actualize, maintain and enhance itself. This seeking is accompanied by emotion, which facilitates the maintenance and enhancement of the organism.

The concept or image of the self arises out of experience and helps to mediate the continual process of adjustment to the environment. Any experience which is inconsistent with the structure of the self image may be perceived as a threat. Maladjustment exists when the individual refuses to admit significant experience into consciousness or, as Rogers says, fails to "symbolize and organize such experience into the Gestalt of the self structure." Threats to the cherished self image bring anxiety and defensive behavior. Under favorable conditions, therapy can produce a reorganization of the value system and reduce the incongruity between experience and the structure of the self. In the normal course of maturation the individual replaces values which have been taken over from others and achieves an inner harmony through greater self-knowledge and acceptance.

This statement by Rogers affords a basis for reconciliation between the two major trends in motivation theory. It also makes a place for the theory of learning emphasized in the earlier behaviorism, but in more dynamic form. We do not learn by passive acceptance of impressions but by the continuous attempt to solve the problems posed by the environment. We make a more skillful use of our means as we learn more about the available instruments. We reshape our ends as we learn more about ourselves. Many achieve individual adjustment without professional counsel because of good family adjustment. A happy household is not only an end in itself but a fortunate setting for informal group therapy.

Marketing management solves its problems by helping consumers to solve their problems. While the good life demands an increasing

¹ C. R. Rogers, *Client-Centered Therapy* (Boston: Houghton Mifflin Co., 1951).

variety of goods, it also draws on the realm of ideas and emotions. With becoming modesty the supplier of goods can recognize that some human problems are beyond his reach. He can well take account of the social and psychological setting in which his products will be used. Half-baked attempts to deal with the problems of disturbed personalities, however, are likely to end up being both poor therapy and poor selling.

A perspective such as that sketched by Rogers can provide some useful guideposts for motivation research. For the rational problem-solver, his own irrational impulses or defense mechanisms are part of the problem. We are learning how to devise experimental procedures which parallel the decision structure of real life problems. Ways can be found to introduce faulty self-knowledge as an element in these experimental designs. Progress in experimentation is a goal which might well attract the exponents of the nondirected interview as well as the advocates of massive sampling surveys. Advances in this type of motivation research should contribute substantially to advertising effectiveness.

Rogers calls his view "phenomenological," a designation shared with other writers such as Snygg and Combs. This means that the environment as perceived by the subject is taken as the behavioral field. Both the world image and the self image of the individual are keys to understanding his behavior. Kenneth Boulding, in his recently published essay "The Image," suggests the term "eiconics" for a science of images cutting across economics, psychology, and the other social sciences. Advertising, which is in the business of creating images, has a stake in the outcome of such a project.

Meanwhile, advertising strategy wisely shows some restraint in applying the results of motivation studies in the present state of the art. The psychoanalytic view in particular is marked by two quite distinct versions from Freud himself, followed by a profusion of free-wheeling speculation by his many disciples. So far the major impact has not been on the advertiser's claims concerning his products, but on the manner of their presentation. A new aura of interest for a product may be created by a fresh copy approach, but the product still has to compete with other instruments recommended for the same purpose.

There is some warrant in clinical lore for calling a spade a symbol of fertility, but it still belongs in the tool shed rather than the boudoir. Every copywriter knows that a man buys suspenders to hold up his trousers and not as a "reaction to castration anxiety." A woman buys a garden hose to water the lawn and not because of the "futility of ure-

thral competition for the female." Possibly we are saved from the solemnities of Freud by the sanity of Rabelais. Any student of the gusty Frenchman will remember a chapter on toilet training that has never been approached in Freudian literature. The five-year-old Gargantua has some remarkable things to say in this chapter about product testing. As for urethral competition, who can forget Gargantua's first visit to Paris and the flood which drowned 260,418 Parisians, not counting women and children. He may have been visiting the Madison Avenue of that day, since his protest concerned the gullibility of the great crowds which gathered to behold any spectacle from "a mule with tinkling bells" to "a blind fiddler in the middle of a crosswalk."

Advertising strategy must take account of both gullibility and gump-tion, of human needs both instrumental and symbolic. In the long run the odds are in favor of a strategy which takes rational problem solving as a fundamental aspect of human behavior. Despite all the quirks and foibles revealed by motivation research, rationality and efficiency are universal goals of the maturing individual. For most products, the long-run advantage probably lies with the kind of advertising appeals that will still make sense to the mature and balanced personalities which most of us are trying to become.

3. TOWARD AN UNDERSTANDING OF HUMAN BEHAVIOR*

Here is a vigorous exposition of the case for the use of psychological techniques in the study of market behavior by a person who is perhaps the most successful of the commercial market psychologists. Notice the manner in which the "conventional statistical approach" is rejected and what roles are assigned to psychological techniques. The defense of small samples, running through the article, is an integral part of this point of view, though the discussion as a whole is on a more general level.

When advertiser, agency, and manufacturer turn to research, they really are interested in finding answers which will enable them to sell more of their particular brand of merchandise or service. Usually the

* Adapted from a talk, "Scientifically Predicting and Understanding Human Behavior," by Ernest Dichter, Institute for Motivational Research, Inc., given at the 1955 University of Illinois Marketing Symposium, and reprinted in *Consumer Behavior and Motivation*, University of Illinois Bureau of Economic and Business Research, Special Bulletin, 1956.

manufacturer knows quite a few of the answers and has made quite a few correct moves. Had he not, he could not now afford the services of an agency or a research organization. It is remarkable, when listening to the discussions of the scientific aspects of marketing research and motivational research, how the greater number of the American industrialists could have had the audacity to introduce new products, change the models of their cars, develop new and successful radio and TV programs—in short, develop American industry at a most amazing rate merely on the basis of ingenuity, creativity, and intelligent hunches.

When the manufacturer turns to research, he does so because he looks upon it as one more tool to help him in his daily task of making decisions. Particularly, when he is attempting to reach and influence consumers, whether housewives or purchasing agents, he wants additional information about their behavior, which will permit him to make additional correct moves. He seeks information which will tell him:

1. Who buys his product and who does not.
2. Who can be sold his product and who cannot.
3. Why people do or do not buy his product.
4. How he can reach more customers with his message most effectively and most economically.
5. What is the future of the market prediction of trends.

For many years the only tool available to the manufacturer was conventional statistical research. The advertiser has gradually come to realize that by using this tool alone he has received only part of the information which he requires to make an intelligent and scientifically based decision. His situation may be compared to that of the medical practitioner who wants to cure a stomach ailment and who has available only an accurate and detailed report on the number of sufferers from stomach ailments in this country, their age, income, diet, and marital status. Much of this information would be very interesting, but only in the most remote fashion would it help the physician to cure stomach ailments.

When the advertiser comes to the market researcher he is interested in finding data which permit him to act, and not academic analyses of the status quo. What does the commercial practitioner need for action data? He needs data of a dynamic nature which gives him an understanding and scientific proof of the real causes of human behavior in the market place.

This, then, is the real and inescapable hurdle which modern market research will have to overcome. No amount of vituperation and emotionality, no attempts to label as unscientific what does not fit in with

conventional nomenclature or may not be expressed in decimal points, can replace a calm and truly scientific analysis of this real problem.

If we can, within the ranks of the broad family of researchers, begin to see that different problems require different research solutions, and that our problems, as researchers, *are* different in many crucial ways, then we shall be in a position at last to unite our band around a common scientific philosophy which nevertheless recognizes the utility of different approaches and tools.

No one involved in contemporary business research would deny the desirability of such a unifying development, least of all the consumers of that research, the businessmen themselves. The accusations and counteraccusations currently rampant in research circles represent more than the healthy, "competitive" claims and counterclaims of robust research organizations. Instead, they are signs of a dangerous confusion and unease—an unease which may very well be communicated to the ultimate consumer who simply will not know which (if any!) of these techniques to choose or reject.

The more we do, therefore, within our own family of researchers to dispel these confusions and doubts among our consumers and ourselves, the more we advance our profession as a whole. It is true that the only kind of research "worth doing" is research that meets the most rigorous canons of scientific accuracy and honesty. But it is also true that science in general, and statistics as its handmaiden, offers many different tools for the solution of the different problems of business research.

DESCRIPTIVE AND DIAGNOSTIC RESEARCH

Business generally has a choice of two types of research: descriptive research, which tells how many did what; and diagnostic research, which tells why what happened did happen. While there is an interrelationship between the two, it is motivational research that provides the tools for creative predictions of consumers' actions in the future and enables the executive to base his advertising and public relations decisions on the pattern of human motivations revealed in the research study.

Studying human motivations is not unlike Herodotus' problem of studying the reason for the inundation of the Nile. By merely observing a person's behavior it is close to impossible to determine why he does what he does. Herodotus approached his problem by picking out, on the basis of previous knowledge, certain elements which he thought were significant. He felt that the distance covered by the flowing waters, the time at which the inundation began, the time at which the overflow reached its maximum, and the fact that there were no winds or breezes

at the river surface were all interesting phenomena, although he did not know what they had to do with each other. In this form they were all meaningless facts, not susceptible of interpretation. Why did he pick these facts rather than others? The answer is that he was familiar with certain theories dealing with the behavior of rivers. It was his familiarity with these theories that made him look to factors like wind, snow-fall, or evaporation, rather than to the quantity of water or the chemical composition of the water in the Nile.

In the field of human motivations, we approach problems with certain general theories about why people behave as they do. For example, we believe that people are more concerned with their own egos than with other people; that most people suffer from a degree of insecurity and have as one of their main goals in life self-protection against dangers and anxiety. Applying this theory to a specific problem, we develop a hypothesis.

Cohen and Nagel¹ state:

We cannot take a single step forward in any inquiry unless we begin with a suggested explanation or solution of the difficulty which originated it. Such tentative explanations as suggested to us by something in the subject matter and by our previous knowledge, when they are formulated as propositions, are called hypotheses. The function of a hypothesis is to direct our search for the order among facts. The suggestions formulated in the hypotheses may be the solutions to the problem.

THE PROBLEM OF CAUSALITY

From one point of view, motivational research deals with the difficult scientific problem of causation. In the conventional market research study, causation can seldom be established, partly because conventional market research must deal with isolated and principally numerical phenomena, and secondly, because most practitioners of marketing research are not trained to establish psychological relationships between data and behavior. When we say that motivational research is "why" research, we are saying precisely that the gap between the act and the psychometric evidence has been bridged. We are enabled, therefore, to say that this takes place because these things have happened. Causality does not imply a simple cause-and-effect relationship. As a matter of fact, it is because motivational research deals with the totality of factors influencing consumer behavior that we can establish causal relationships. It is the insight brought by this type of study to the

¹ M. R. Cohen and E. Nagel, *An Introduction to Logic and Scientific Method* (New York: Harcourt, Brace and Co., 1942).

businessman or advertising executive which permits him to use motivational research and immediately and effectively translate it. We do not attempt to provide in our studies a table of the frequency with which a particular causal relationship exists in the population as a whole. What we have attempted to establish is the significance of the causal relationship, and not necessarily the frequency of its occurrence. In establishing this causal relationship we bring to bear on the problem, not only findings in the individual study, but also the findings in all of our studies, plus the accumulated knowledge developed by the social sciences.

MULTIPLE FACTORS

It should be emphasized that relationship between cause and effect cannot be attributed to a single, simple, motivational factor. The real problem of motivational research is to analyze, and identify, the variety of interrelated factors which culminate in a particular action. This cannot be done by a simple analysis of the frequency of behavior or the frequency of a typical stylized response to a direct question. The very complexity of human motivations requires the broadest approach to their analysis combined with a rigorous multiple validation which defines and isolates the individual factors.

MOTIVATIONAL RESEARCH IS QUASI-EXPERIMENTAL

Experiments in the laboratory sciences are designed to establish the relationship between variables in a controlled and usually quite isolated corner of the universe. Usually this corner (the laboratory plus the experimental materials) is most atypical of the universe as a whole. For a physicist wishing to test the magnetism of certain types of steel, there is no question of representativeness: there is no "sampling problem." Why?

Well, for one thing, a "random sample of pieces of steel" would be a most cluttered and disorderly collection of odds and ends, and utterly useless for his purposes, not to mention the fact that it would be fabulously expensive to obtain. What *does* he want, then? He wants good examples, as free of impurities as possible, of the several types of steel about which he has erected hypotheses. How does he get them? He defines his types, and then goes out and gets "good" (that is to say, measurably not atypical) samples of each type.

Now suppose that one of these types is just as frequent in the real world as all of the other types put together. Does he arrange his sample so that half of the steel pieces are of this type? He does not; it would be

foolish, and it would certainly blur his results. For the scientific fact is that the rare types are probably more important in his analysis than the common ones.

In applying this general rationale to motivational research, we may, for example, use anywhere from forty to fifty case histories taken from our consumer motivation panel, where we know not only how the individuals behave, but what their histories have been within the normal realistic family and community background. These preliminary steps help to produce a series of meaningful hypotheses which are weighed and checked against all the accumulated knowledge in the field of human motivations. Not unlike a good diagnostician or a detective with many years of experience, we can set forth tentative explanations of why a particular event has taken place.

PSYCHOLOGICAL VARIABLES

In defining our approach to a problem, we seek to establish the major psychological variables which may influence the findings. In a sense, this is an aspect of the development of hypotheses. What we are searching for are psychological and personality elements which may have a dynamic effect on consumers' attitudes toward a product. Such variations will often cut across income, educational, and professional lines. In doing so, we not only take into account such factors as use, nonuse, and frequency of use, but all of the psychological variations which affect the problem.

For example, in a recent study, we established the fact that the consumer may be divided into two groups in terms of his attitudes toward milk. One of these groups was the "milk-involved;" the other was "milk noninvolved." These were two of the important psychological variables which had a definite relationship to attitudes and feelings toward evaporated milk. It was by developing this psychological variation that our study findings gained real significance. The analysis of these variables is of extreme importance in understanding the real responses of the population to particular situations. Such factors as adventurousness, insecurity, prejudice, self-indulgence, and so on are often of more significance in determining market responses than such objective factors as income, age, and so forth.

MULTIPLE VALIDATION

The second step of our task is to verify or refute, exclude or reformulate these tentative explanations, and to study them and their varied ramifications. A number of avenues are open for implementing this step, the rigorous statistical testing of our hypotheses. We can conduct

depth interviews and do observational studies noting the actual behavior of people. We set up trigger systems of circumstantial evidence, stating that if Hypothesis A is correct, Events 1, 2, 3, 4, and 5 should occur in a particular manner. Then we go out into the field to see if these events do indeed occur and conclude thereby with adequate scientific justification whether Hypothesis A was correct or incorrect.

This second step, the verification or refutation of our hypotheses based on approximately 200 or 250 individual histories, still does not lead to the same kind of variety and numerical accuracy, as for example, ability to distinguish between a brand-recognition index of 68 per cent or 73 per cent, that 2,000 or 5,000 interviews would do. What we have contributed, however, is a really thorough understanding of a basic motivational pattern among a group of people large enough to indicate that the pattern is significant and lends itself to practical applications. In finding that 80 mothers out of 100 reveal, in multiple waves and multiple tests, when talking about food for their babies, that they are as concerned about their own convenience as they are about the nutritional value of the food, we have a finding valid enough to permit any practitioner in advertising or public relations to take advantage of it and act accordingly.

Using various research techniques to test a hypothesis permits multiple and structural validation. A central theory may have a number of subsidiary hypotheses, each of which points to a different corner of the phenomenal world. So, armed with a number of small samples, each of the most efficient size possible, we go out and look in these corners. We use field tests, projective tests, interviewer-administered questionnaires, all kinds of measures—each applied to its own distinct sample of respondents. Then, if supporting tendencies are found in each separate pile of results, we have something much more reliable, scientifically, than one massive sample could have produced.

BASIC MOTIVATIONAL RESEARCH TOOLS

Motivational research employs a gamut of tools, from simple word-association tests to complex and detailed projective techniques. In all, the basic aim of these various procedures is to develop sufficient information to distinguish rationalization from reason.

The two most important techniques employed by our organization are the depth interview and the projective technique. The depth interview has been much discussed.² It is employed to elicit the freest possi-

² For a clear-cut explanation of the basic approach taken by the depth interview, and the ways in which our organization uses it, see Robert K. Merton, *Mass Persuasion* (New York: Harper and Brothers, 1946).

ble associations on the part of the respondent. The respondent himself, for the most part, determines the direction, the order, and the content of the interview. The interviewer must be skilled at developing rapport, inducing the respondent to express himself, and inserting delicate probes, where necessary, to encourage fuller discussion. Scientifically speaking, the depth interview begins at the point where the open-end question used by some conventional research organizations ends. It avoids the interruption of rapport involved in following a set procedure. Instead, by giving the respondent an opportunity to talk about himself, listening sympathetically, and encouraging further self-exploration, the closest kind of rapport can be developed in the course of depth interviewing procedure. The mass of information accumulated by the depth interviewing technique is analyzed for the purpose of determining the *meaning* of the consumer's behavior rather than relying strictly upon her own explanation. In this area, we bring to bear the specialization of the members of our staff in the various branches of the social sciences.

The projective technique is basically a method of inducing the respondent to talk about herself in a disguised form. Where, for example, we may be encountering motivational patterns which are socially unacceptable, the projective technique, in which the respondent interprets a deliberately vague story, photograph, drawing, or design, permits her to attribute her own motives, without feelings of guilt, to some third party. It also permits her to express feelings about the external, which she may not know about herself, yet which are extremely revealing of her own attitudes.

Whenever it is desired and practical, there is no reason why such a pattern cannot be tested on the basis of 5,000 cases. In practice, however, at the end of a study a client will much more frequently accept our recommendation, if it provides him with new insight into his sales and advertising problems.

WHEN ARE MOTIVATIONAL TECHNIQUES SCIENTIFICALLY CORRECT?

The use of motivational research techniques is not just a matter of voluntary decision. There are definite circumstances when motivational research techniques are the *only* ones which will produce meaningful results. The use of conventional market research techniques, under these specified circumstances, may be unscientific and misleading.

What are the circumstances in which motivational techniques and projective techniques are indispensable?

1. Whenever the data being sought may not be present at the rational or conscious level.

2. Whenever we are dealing with psychological mechanisms and not with simple static cause-and-effect relationships.

3. Whenever the respondent has a chance to produce interference consciously or unconsciously between the time he understands the question and the time he answers it.

Human motivations are frequently unconscious and represent intertwined and complicated mechanisms. Modern psychiatry and all other social sciences could not exist without the acceptance of this fact. To explain this position, the really trained modern social scientist with thorough clinical knowledge finds himself in the same position as the researcher who attempts to convince the patient who "is sure he knows what causes his stomach disorder" of the existence of invisible microbes or of mysterious cells which suddenly start growing beyond their intended size and function. It is as difficult for some of us to accept the existence of the often exasperatingly irrational aspects of modern psychoanalytic and social psychological concepts as it may be for the unsophisticated layman to accept the doctor's statement that the headache you complain about really comes from an infection between your toes.

Those criticisms which contend that most social science concepts are contradictory and that no two psychologists can agree between themselves about the interpretation of an observation are of a similar nature. It is true that three doctors might interpret an X ray differently. This does not prove that there are not ever-growing areas of agreement in the medical and biological sciences and that we should stop sponsoring a cancer institute.

In a recent book, *Psychoanalysis as Science*,³ a group of scientists set it as their task to establish whether or not psychoanalytic explanations of human behavior could be proven by rigorous experiments. A reading of these experiments is highly recommended.

If the existence of unconscious motivations and their basically dynamic and complex nature is accepted, we have come much closer to answering the question of which research techniques are correct, necessary, and desirable and which ones are not.

It is my view, based upon studies and experiments, as well as the mass of psychological scientific literature, that whenever we deal with human motivations, direct questions are not only inadequate, but unscientific and therefore are to be rejected.

The systematic elements of my argument are the following:

1. *Each individual represents a motivational universe.* In order to be able to record scientifically how a population is motivated in respect to

³ E. Pumpian-Mindlin, ed., *Psychoanalysis as Science* (Stanford: Stanford University Press, 1952).

a specific phenomenon, we must first fully understand why each individual, as a member of this group, behaves in the way he does. Unless we have arrived at such understanding, no meaningful statement about the motivations of the group can be made. In other words, we have to be clear at the end of a motivational interview why Mrs. Brown buys a particular brand of coffee. This evidence may include a wide variety of contributing factors.

2. *Direct questioning erroneously presupposes objective insight.* "Know thyself" is one of the oldest goals set for the individual. We have come closer to it, but we are still far from achieving this goal. Whenever we are asked why we did what we did, we must analyze our actions and give a correct and insightful answer, often within the limit of a few minutes. Even if no unconscious factors are at work, we are often incapable of choosing, from among the various conflicting motivations of which we may be aware, the one which influenced the resulting action.

3. *Rational explanations and rationalizations.* The human personality does not represent a passive observational field. You cannot simply consider a question as an objective means of discovering what goes on in the consumer's mind. Whenever a question is being asked, an active interpersonal process is instituted. An interviewer or a printed form challenges me, as an objective outsider, to give an account of myself.

Psychology has demonstrated that there are several permanent distortion factors which interfere with the objective observation of the motivational field. The most important one is our desire to appear rational to ourselves and to others. When confronted with an investigation of our motives, we first search actively for rational explanations. The danger is great, however, that this desire to act rationally results in a rationalized answer, a pseudo-rational cause for our behavior.

THE CONCEPT OF "UNIFIED" RESEARCH

Of course motivational research is not a panacea. There is good and bad motivational market research, just as there is good and bad conventional market research. What is needed is a unified concept in research, dealing with the behavior and motives of human beings.

When the advertiser wants to be able to predict the outcome of his actions and asks for help, he is asking for motivational answers. If the advertiser is satisfied with knowing that most people will drink tea when they are ill, that more people have one car than two cars, use one brand of cereal, or listen to a particular radio or TV show, then the techniques utilized to approximate census figures should be fully satis-

factory to him. If the advertiser, however, is not satisfied with finding where his sales are good and where they are bad, but wants to bring about changes, wants to sell a second car to the family, make them buy his brand of cereal, and listen to his TV show, he is invading the field of the social sciences and particularly the field of the psychological disciplines, whether he realizes it or not.

The moment you are dealing with human beings, their opinions and their emotions, you are in the field of psychology. It is only too true that the social sciences are far from perfect. But this is a complaint that all of us have to live with whether we employ conventional market research or motivational market research.

There is no doubt that motivational market research needs discipline. But it has to be the discipline appropriate to its specific nature as a science. To insist that, because you have to research 2,000 people to know how many people have stomach disorders at a given time, any scientifically proven explanation of the real causes of stomach disorder based on experimentation and an entirely different set of inductive and deductive inferences is invalid, is in itself proof of lack of scientific discipline of a much wider consequence.

Prediction of consumer behavior necessitates first understanding of his behavior today. As an individual reacting, and at the same time influencing the world within which he lives, the motivational researcher is at the same time a participant and a shaper of the future world.

4. HOW PSYCHIATRIC METHODS CAN BE APPLIED TO MARKET RESEARCH*

Here is another strong espousal of the value of psychological (psychiatric, sic) techniques in marketing studies. In contrast to the general theoretical approach taken by Dichter in the foregoing pages, this article takes a more specific approach and presents six types of techniques adapted from psychology for solving marketing problems. In each case, the illustrations are clear and to the point.

Good research, like good journalism, answers the *who, what, when, where* and *why* of a problem. The standard survey method, with its

* Adapted from an article by James M. Vicary, research consultant, "How Psychiatric Methods Can Be Applied to Market Research," *Printers' Ink*, Vol. 235 (May 11, 1951), pp. 39 ff.

emphasis on nose-counting, starts off strong on the *who* and *what*, weakens a bit on the *when* and *where* and frequently fizzles on the *why*.

People can usually be stimulated to tell what they *think* are the reasons for their behavior or preferences. All too often their rationalizations do not stand up. They may say they brush their teeth to prevent decay, but actually they brush their teeth most frequently before breakfast; they may favor anti-chain store legislation, but most of their shopping is at the local A & P; or as happened during the last war, they may favor strict rationing laws, but actually engage in black-market buying.

There are ways to get behind these rationalizations. Rather than ask very precise and rigid questions, the approach is indirect. Responses from subjects are induced with purposely vague questions. The respondent is obliged to pour meaning into only partially defined situations, and this forces him back upon his own inner psychological resources. Here's how you can apply six approaches to business research developed in the field of psychiatry.

WORD ASSOCIATION METHODS

Word association is probably the best known of these methods. The respondent is asked to give the first word that comes to mind for each of a list of unrelated words. Given the word "bread," the first word that comes to his mind is "butter." Despite its simple operation, it is a highly flexible method. It is a semantic test revealing the content of people's basic associations. Besides that, analysis of the time it takes for the respondent to give a response and his ability to give the same response quickly during a second test reveals his stability of association and the extent of his emotional reaction to each symbol.

By all odds, the most obvious application of this method is screening a long list of names proposed for a new product or model. For a new beer to be nationally distributed, it was found that people could not easily and readily associate beer with foreign-sounding names. In some instances they confused this type of name with cheeses and other products.

New words to describe old products can be screened in a similar manner. Coined words for trade-mark names in particular are tricky. They are much sought after by advertisers because they can be legally protected, but their novelty of expression may create difficulties. Take a term like "Solium," which has been heavily advertised. A totally new coined word for the same type of ingredient produced two and a half times as many appropriate associations as "Solium." An important group

of respondents confused "Solium" with drug or cosmetic products. Words like "complaint," "cooperate," and "voluntary" have been shown to produce deep emotional disturbance with sizable groups of the general population and are dangerous to use in most contexts.

Controlled word association is a variation in which the respondent is asked to give only a certain class of word response. For example, a list of products is read to him one at a time, and for each he gives a brand name. Professor Houghton's *want-association* and Dr. Link's *triple associates method* are examples of the controlled approach.

Another example is the study made recently by Joe Belden, Joe Belden and Associates, Austin, Texas, for an advertising agency. The objective was to explore the use of word association tests to measure the impact of advertising by the agency's clients.

A cross section of 250 adults was asked to name the first business concern that came to mind when confronted with certain words, such as hotel, bread, bank, house paint, and dairy. The words were rearranged and the same question asked again. The same procedure was used with a list of the agency's clients in those businesses. As a result, Mr. Belden secured a statistical view of associations of products and services with certain advertisers, and vice versa. The result was a view in depth of the relative impact of each of the advertisers involved.

Successive word association consists of asking not just for the first word that comes to mind, but also for a long chain of associations. To the word "meat" a respondent might give, "potatoes," "steak," "food," "bread," "fish," "red," "eat," "beef," "pork," and so forth. The method is useful in studying ways to change associations by substituting one symbol for another. It is also a means of quickly canvassing a person's associations on a particular subject. For example, before making up a questionnaire on soup or some other similar product, the researcher might do well to write a long succession of associations in order that he will be fully aware of all the allied subjects.

COMPLETION METHODS

The *sentence completion* method is the most familiar of the completion group of methods. In this type of questioning a word or phrase is given, such as "When I have a headache", or "Meat" The respondent is asked to add any words that come to his mind to make a complete sentence. In the first item a respondent might say, "I usually use Bayer Aspirin."

It can be seen that this method provides greater context and more information for the test item than the word association method. In com-

bination with the word association method, the sentence completion approach is capable of very sophisticated use. In one instance where people would not openly admit their fear and annoyance about increased airport activities near their homes, a combination sentence completion and word association interview was successful in showing the character and extent of this hidden fear. The fear and annoyance was far more extensive than standard interviewing methods had indicated.

Wherever a respondent's answers might be embarrassing to him, as in drug or personal products or during times of crisis like strikes or failure of accepted practices, it is relatively easy to study these areas with such methods. As an added element, your respondents and even your interviewers need not know the purpose of the study, since rather complete masking of test items is necessary for both methods. *Story completions* have also been used in much the same manner.

PICTURE AND VISUAL METHODS

Both the *Rorschach* and *thematic apperception* tests employ pictorial material. In the apperception test, subjects are asked to tell a story about a picture. Most of the material revealed in this test is either directly or indirectly autobiographical, although respondents are not usually aware of the fact. These two methods in their orthodox use require technical training, and readers are referred to psychiatric manuals for an adequate discussion of their use.

Simple variations of the thematic apperception test have been employed in commercial research. Probably the most striking is the use of the *comic strip sequence*, in which some of the balloons containing conversation of characters are left blank for the respondent to fill in.

ROLE-PLAYING AND SITUATIONAL METHODS

The forerunner of the role-playing and situational type of test is the *psychodrama*, originated by Dr. J. L. Moreno. In personnel research in particular this method has been applied successfully. The respondent acts out a significant event in his life or plays the role of one of his superiors or subordinates. When an antagonistic employee, for example, acts out his impression of the boss's job, the reason why he is antagonistic may be revealed.

A *personification* method has been devised in which a respondent acts out the role of some nonliving thing, such as a magazine, product, or company with which he is very familiar. This method was used to find out whether the public understood the editorial policies currently in force by large circulation magazines.

In another approach, respondents are asked to imagine in detail the kind of person who has bought a number of times at a store. One of the items in the list is under study; the others are used as a mask. The same list is used on a matched sample, except that the test item is changed to another product or brand for comparison with the test item.

Mason Haire of the University of California described in an article in *The Journal of Marketing* for April 1950¹ how motivations for not buying Nescafé instant coffee were uncovered by using this approach.

CHAIN INTERVIEW METHODS

Since most people's information comes to them by word of mouth, it is important that we have a method of testing this natural chain reaction of a message. In the *chain interview* method we duplicate the children's game of telling a story around a circle in which each child whispers the tale to the next until a distorted version comes back to the first child.

The research method incorporates electronic recording of all the testimony in the chain so that each person's contributions in the chain can be studied. Each respondent is urged to repeat the test message as exactly as possible, adding any bona fide information to it which he considers important.

The experience of several different investigators shows that names of places, people and things normally drop out of the chain very rapidly. The possibility of a break in the chain of identification for a brand or company name is very important. The associational strength of the original name and the context in a message are both crucial in resisting this natural deterioration as people pass on the story. This test can show how far word-of-mouth advertising will carry your slogan, brand-name, price, and so forth.

Perhaps more dramatic is the use of this method in picking up in advance of release any possible boomerangs in a message. In one case the announcement of a retirement fund for the elderly executives of a firm was found by this method to be capable of being twisted into a bad connotation by customers. It was said at the end of several chains that the graybeards were to be thrown out of the company pronto! The reader will no doubt recall actual examples where sudden or repeated price reductions, entirely justified on an economic basis, have been very damaging to the product's reputation or that of the manufacturer. Something like this happened with the ball-point pen. A price reduction, announced without proper explanation, may start an idea rolling that cannot be foreseen without a pretest of this kind.

¹ This article is contained on pp. 93-103 of this volume.—ed.

DEPTH AND CONVERSATIONAL INTERVIEW METHODS

The so-called "depth interview" method is probably the most popular way today of obtaining qualitative information. Interpretation of attitudes by this method is extremely difficult, primarily because it is not a systematic method. As a pretest device before conducting the standard survey, it has been most valuable in obtaining more than an arm-chair view of a problem to be studied.

Most of the methods outlined previously are not costly. They do not run to the volume of the nose-counting type of operation. Many of the problems uncovered by these studies, however, can be tested and given some validation by the survey method. Here the costly hit-and-miss method of speculative surveying is avoided, because the survey is limited to known problem areas.

There should be more use of these methods in market and public relations research. Too many researchers have been trapped in the assumption that mere statistical manipulation is a substitute for thinking. These methods force a more careful analysis of the reason behind consumer actions and reactions. They can be used to uncover the reason-why of success and failure in the field of marketing.

5. A CASE STUDY: WHAT AUTOMOBILES MEAN TO AMERICANS*

The manner in which a psychological approach can provide new insights into marketing strategy is illustrated in an excellent fashion by this study on why people buy particular automobiles. Although the emphasis is on the personality of the buyer, it will be seen that economic considerations are not ignored, and it is interesting to note how the psychological factors related to marketing behavior are fitted into an economic framework.

Year after year the automobile moves farther out to a center-stage position in the minds of most Americans as virtually the most exciting thing in their lives. The cost of purchase and upkeep rank second and third in the average family's budget, yet no one ever raises an

* Adapted from a talk, "Automobiles: What They Mean to Americans," by Pierre D. Martineau, Director of Research and Marketing, *Chicago Tribune*, given at a Seminar in Motivation Research, University of Michigan, 1954, and later printed as Chapter 6 in Pierre D. Martineau, *Motivation in Advertising* (New York: McGraw-Hill Book Co., Inc., 1957).

eyebrow about the wisdom of such an expenditure. In our national thinking, we assume automobile ownership is as vitally necessary to decent living as indoor plumbing and adequate shelter. It actually startles us occasionally to encounter middle class families without a car.

The sales curves for the industry testify, moreover, that this desirability is becoming more intense all the time, to the place where the automobile seriously interferes with the prosperity of any number of unsuspecting industries. Whether they know it or not, almost all retailers and manufacturers selling the consumer are competing with this American passion for dazzling new cars. When the family reaches out for more and more fascinating extras, sporty convertibles and station wagons which are more expensive than the less-exciting conventional models, then it has less to spend on other things like jewelry, apparel, entertainment, and furniture.

Why is this so? Why does the car mean so much? What intense gratifications does it provide to create this remarkable desirability? Why is it assuming greater desirability?

Of course it is a means of transportation. This is perfectly obvious, and it is a practical reason that no one can minimize. But it represents much more than transportation. If that is all it stood for, we would buy inexpensive cars and drive them for 150,000 miles. The mere factor of transportation sheds no light whatever on the public mouth-watering for brilliant colors, for two-tones and three-tones, for rear fender aesthetics and chrome side-strips and lush interiors. Nearly all cars today are mechanical achievements, yet the individual decides he wants a Ford or a Buick instead of some other car. Why?

The potentialities offered by competent qualitative research for revealing insights on the dynamics of behavior, and how it opens up new directions for thinking, are brought out with reference to an automobile study conducted for us by Social Research. This study touched on the following specific facets of automobile ownership:

1. The car is pretty close to being the most desirable product in our economy, and it is worth examining the motive elements in this desirability just to see how any product can climb so high.
2. Although it is essentially a complicated piece of machinery, the car has attracted to itself a wealth of other meanings far beyond its functional activities.
3. The car has embedded itself very successfully into the currents of change in our national system of values.
4. This study throws significant light on the whole process of decision making in any field, although it is more obvious in the case of the automobile.
5. There are some highlights on the basic objectives of advertising itself.

6. This study documents an entirely new dimension in advertising and marketing—how a product in reality is seeking to sell itself to various human personality types, each of which may have different goals and different self-conceptions.

The findings for this study are based on both structured and unstructured interviewing. A number of attitude and opinion studies were used effectively for their directional inferential values. Besides a number of verbal projectives included in the free interviewing, modifications of three standard projective techniques were employed—the Thematic Apperception Test, Sentence Completions, and Role Playing.

While all of the interviewing was done in Metropolitan Chicago, experience with this and other studies would indicate that the findings are pretty generally typical of urban American attitudes, which are created in large part by the products themselves and their national advertising.

The sample size (350) was fairly large for qualitative interviewing, and was deliberately weighted to include two-thirds men and one-third women. Five-sixths of the subjects were car-owners. Instead of the classical index of income distribution, the sample was stratified by social class as a much more precise description of buying behavior and general attitudes.

It should be stressed that the value of these studies is first and foremost directional. Quantifications are not attempted except in the roughest sort of fashion. The small sample was broken into subsamples to permit using more qualitative techniques, all of which have acceptance in the social sciences, and which have created a piece of diagnostic research going much farther as explanation of consumer behavior than any of the rigidly quantified techniques.

THE AUTOMOBILE AS A SYMBOL

Occasionally some penetrating mind calls attention to the fact that the American market consists of a multitude of markets—different age groups, the two sexes, racial and religious groups, social classes, urban and rural distinctions, income and occupational levels, geographic viewpoints, and probably many other factors generating different wants and different standards of taste.

Personality is certainly not the least important. Ordinary criteria for separating buyers often give no explanation for product successes and failures. Buyers and nonbuyers may be indistinguishable except on a personality basis. Thus, extensive studies of sales abilities of

department store clerks indicate that the crucial factor is one of basic personality components. Another study of unsuccessful business executives established that their failure was attributable not to knowledge lacks but to personality problems.

Our automobile study very clearly shows that the market for a Plymouth or a Pontiac actually is not all the people in a certain income group or a certain locality, but rather distinctive personality types. Neither does this mean, however, that personality is the only factor, but it is a dimension that should not be overlooked. Many a product can carve out its own segment of the market by creating a specific appeal to a definite personality type, instead of trying to be all things to all people. Smirnoff Vodka, Parliament Cigarettes, and Schweppes Tonic Water are concrete illustrations of this achievement, becoming meaningful and important to a definite group, small in per cent but large enough in number to constitute a profitable market.

The practical, social and personal meanings which the automobile holds are grouped around five central ideas.

First, of course, the car is a mechanical object with important practical uses. The car as a means of transportation satisfies many practical goals and needs which are motives in themselves. These are the sensible reasons why people want cars. We have become such a physically mobile people that transportation is literally a key factor in our physical and social life. This element of transportation is so obvious, however, that it is unnecessary to belabor it.

Much the same is true about the technical features. The car is certainly an extremely complex machine with a fascination for our national passion for technology and gadgetry. Nevertheless, it is highly important to indicate that the number of persons with overwhelming mechanical interest is definitely in the minority. Cars have become so complex and so mechanically perfect that the average driver understands very little of the intricacies of operation.

In this study, people were asked to tell a story about a picture of a car with the hood raised and a man peering at it. There were two women in the back seat. The stories all identified the situation as a breakdown on the highway. But a general helplessness was indicated about how the man would fix his car. He should look for a garage or a service station, and not tinker with it himself. The average motorist today rarely if ever raises the hood on his own car. He has only the barest notion of car mechanics.

Of course we like to talk about the spectacular new engineering triumphs, particularly the well-advertised features. They are conversa-

tion pieces on the American scene. The individual betrays that he is out of touch with things if he cannot at least identify such wizardry as power steering and hydramatic drives. But our interest is much more in the gimmicks and latest innovations, not the complexities. And we are far more interested in what they will do rather than how they operate.

A second group of meanings revolves around the automobile as a sizable investment with an important place in family economics. The individual has to think about it often and seriously. It never gets pushed out of mind by other things.

It is important to point out how these economic factors work so they won't muddy other ideas which parallel them. Cost is only one of several considerations, and usually is not the determinant in the choice of one make instead of another. All of us are exposed to an avalanche of spectacular bargain claims from less popular cars. But we yawn them off. Unless the make is appealing, we aren't particularly interested.

The automobile retailer probably would disagree with this. He knows in one sense the power of excessive dealing. But today the car dealer does little to sell the individual make. The fortunes of any car over the country, its up-curves and down-curves, are pretty uniform regardless of dealers. When the prospect arrives at the retail showroom, he has already been presold on the make he prefers. Naturally he wants the best offer he can get on that car. But the point is, he has done more thinking, both consciously and unconsciously, about the make than about the deal he might get.

People in this study actually spent more active thought about cost of upkeep and gas mileage than about the price of cars. This was true in any price class. Regardless of income level, every American wants to fancy himself a shrewd and careful manager. So he nods his head owlshly about low gas mileage, and frequently tries to convince himself and his friends that this was the reason for his car choice.

Almost every person I know with a European sports car has tried to justify it by describing how much gasoline he saves. He is obviously screening other motives. His gasoline savings would have to be fantastic to compensate for the difference in car price. Cost of upkeep is not a powerful motivating idea. It is not likely to be the important reason for buying some particular make.

The third general field of social meanings surrounding the car should in large part be apparent: how much the car is involved with people being together—family vacations, driving to work with a friend. Social companionship is a powerful motive, but this factor is quite

obvious. Yet it is important to detail how much the automobile in American life has come to be used as an indicator of the individual's social status, his position, his place in the scheme of things.

Regardless of whether this is utterly illogical, we unconsciously classify every car, and use it as a yardstick for determining how important a person is in our society. When we observe a stranger getting out of his car, our behavior in considerable part is determined by the make and the age of his car. We go to college reunions and assess a classmate's lifetime as a success or failure by taking a look at his car.

Furthermore, the automobile is a portable symbol of status. Only a handful of people can be aware of my bank balance, my investments, my wife's fur coat, my country club membership. But the car is a useful device for carrying my position and accomplishments with me for all to be aware, wherever I go.

Price is one more indicator of a car's status, but not a cause of it. They do match up very closely, however. It is permissible for any American to buy whatever car he chooses. The car is really a prestige object, like a fur coat or a Florida vacation, that anyone can help himself to if he can afford it.

Car selection, nevertheless, is very much influenced by its own status with other cars, and by the status and personal wishes of the buyer. Some people, of course, will be reaching, because the individual's dreams, his hopes, his wishes and his own notions about himself play an extremely vital role in the automobile's meaning.

The average man can buy a classy car and still be wishing. He can enjoy all his day-dreams about wealth, importance, sophistication every time he steps into his car, and yet he hasn't tipped his hand as if he tried moving into a neighborhood too far up the ladder. He can leave all his dreams in his car with no one being the wiser.

The halos of different cars vary considerably, some as chic, sporty, flashy, whereas others are more down-to-earth, less pushy. Some low-priced cars are most acceptable as second cars, others definitely are not.

In the fourth set of meanings, the automobile emerges as a powerful symbol of self-control. It clearly signifies personal mastery and control of basic human impulses. The car is a mass of enormous energy, and because the driver is controlling all this power, he knows the pleasures and fears of handling major drives.

Self-assertiveness is a basic impulse expressed in many motives. Competition, rivalry, the individual's desire for authority and self-importance are manifestations of this urge. The car provides deep emotional stimulation by satisfying this drive through a sense of

power, with the thrill of speeding, by offering many opportunities for showing personal superiorities.

Affiliation is the expression of the need for human companionship, such as sex, friendship, anything that pulls people together. The auto plays an obviously important part in our mating and dating habits.

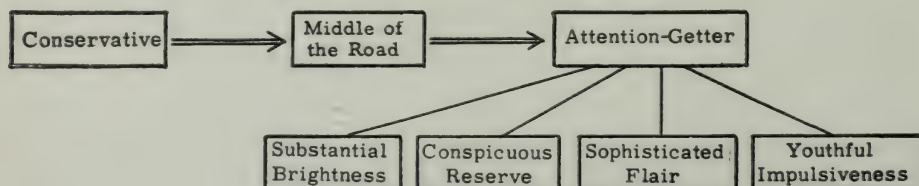
A major part of the car's intense value is that it provides socially acceptable outlets for these basic human drives, while at the same time it is a massive concrete symbol of our self-control, to which we attach a high premium. Any driver senses he is directing potentially powerful forces. Because of this, cars are very much able to satisfy the important desires of Americans wanting to be somebody. Every person wants to be important and effective, in ways that other people recognize. I want to be a good bridge player, a good dialect storyteller, or a stand-out in my work-field. Most children and adult games allow each person in turn to be "it," which provides us even a momentary importance. We crave spoken and unspoken praise, if only from imaginary audiences.

For many reasons and in many ways, having a car gives us the feeling of being somebody. The man who owns a car or the youngster who is even driving a borrowed car feels more important, more substantial, more achieving, more masculine than the man without a car.

The fifth area of meanings has to do with the car as an extremely important avenue of self-expression. The automobile is one of the most clearly understood ways in our society of conveying characteristics and feelings and motives that typify a particular individual. The car tells what we want to be as persons—or what we think we are.

INFLUENCE OF PERSONALITY

Because these personality differences in people have so much to do with the way they express themselves in car selection, this study indicated some of the broad lines of what essentially different people want to be. These dimensions were drawn to cover a normal range, instead of trying to examine clinical types. People were construed to exist along this range:



In general terms, the wish to be conservative or reserved is recognizable in the solid citizen who emphasizes conformity, wants to be considered reliable and stable, draws back from anything conspicuous or drastically new, and who doesn't want to experiment with fads and fancies. In automobiles, the conservative tastes very definitely weight his preferences in makes and models and car qualities.

The conservative prefers the less flamboyant colors, and has been the last to accept lighter colors and two tones. He hangs back on each new innovation, he walks a middle road on chrome and extras, he wants a neat looking new car rather than a high style or extreme design. Certainly he thinks his new car is exciting and shining, but he definitely doesn't want a jet plane or an all-white car. In all of his buying, he stresses money values—in this case real economy, trade-in, and dependability. These factors are more important to him than appearance or power.

The next broad group of personality types are people who emphasize the wish to be sociable, who don't want to be extreme conservatives or too advanced. They are the ones anxious to be up to date, to be popular, to do the fashionable things. If many people are moving toward glamor and sportiness in their tastes, this group wants to be on the bandwagon. Consequently, this attitude is the most suggestible, and the easiest to influence. It is essential to establishing a new fashion.

In cars these people will generally prefer what is currently thought smart without being extreme. They want new styling, and they worry less about mechanics and more about appearance than the conservative.

The third personality type stresses the wish for attention. This is a perfectly human trait which in some degree is in everybody, but in this group becomes very important. These people especially want to be different and outstanding; they enjoy being talked about. Consequently, they are the innovators and the style setters just because they are looking for the newest things. They don't hesitate to try some novelty.

All these people would like to be noticed, but how they express it will vary. There are four main directions of this attention-getting wish:

1. Those who go in for substantial brightness: big, ornate costliness in cars; loud, impressive displays; everything the best and most expensive. They especially want the widely recognized high status symbols.

This could be the car owner who wants to cut an important figure by having everything the best and most expensive, who is anxious

for others to realize he has arrived. But this also includes many people who feel denied in certain areas of living, like Negroes who are overly anxious to prove they are equally as good as anyone else by driving big, shining cars.

2. Conspicuous reserve characterizes a type that wants others to be aware of their status, but at the same time they would like to express modesty. They usually go in for dignified display or deliberate downgrading. This is a familiar technique among people perfectly certain of their higher social standing. They assume indifference to position by purposely buying less expensive cars than might be expected—dilapidated station wagons, old cars, lower-priced lines. The upper middle class is in too precarious a struggle for status to take any such chances.

3. Sophisticated flair describes young intellectuals, sophisticates, professionals, creative people and others who lean toward *avante garde* living. They especially prize individuality in car styling and colors, they are seeking smartness rather than gaudiness, they want bright colors just as they shun drabness. Being in the forefront of style is extremely important; consequently, they will try anything—foreign cars, white convertibles, estate wagons.

They are the people who introduce pizza, turtle-neck sweaters, moccasins, and exotic-sounding vodka mixtures. They think of themselves as being very individualistic, considerably ahead of the generality of persons in taste, intellect, and colorful personality.

4. Youthful impulsiveness is characteristic not only of youngsters sighing for gadgets and hot rods, but it also includes the eternal sophomores of any age who express their youthful yearnings through automobiles. The sixty-year-old driving a convertible with the top down, wearing his trench coat, a crew cut and no hat, is very much expressing his eternal youthfulness. For them a car is a means for demonstrating their daring and their youth by the way it looks and how they drive it.

This is only one very elemental scheme for classifying personality, and it should be made clear that most persons are probably a mixture of many of these characteristics. Furthermore, all of this has been vastly oversimplified, but it was purposely done in order to introduce the factor of personality.

INTERACTION OF BUYER AND PRODUCT

On the basis of this and many other studies, it is apparent that any buying process is an interaction between the personality of the

individual and the so-called personality of the product itself. They are not the same thing, of course. Product personality can be called its character, its reputation, its image. Essentially, this is the whole set of attitudes that people hold toward it.

A car or any product has to accomplish certain purposes, but how I decide which particular auto can best fulfill the job becomes a personal expression of me. In any price class there is a wide range of choice. The car or the brand I pick expresses what I think I am—or what I want to be.

To decide anything as a choice is actually to express the self. Any person who makes a decision among alternative choices is actually stating, "This is the kind of individual I want to be. This is me." What is a self but a permanent mode of selection?

It is important to realize that cars in exactly the same price class can be appealing to very different people. For instance, the lines in the fine car field are actually appealing to quite different markets.

Cadillac has become the outstanding status car, with a wonderful set of favorable attitudes haloed around it. It is considered the most luxurious, the finest engineered, and the most impressive car. It has been transformed in recent years from the high society car to the symbol for the man of achievement, the realization of our American dream of business success.

Almost more than any other symbol, the Cadillac signifies the end of the rainbow for the man who has with his two hands carved out his own road to financial achievement. Curiously, in many companies the president drives another, less-expensive car because he has had his position and money for a generation or so. He doesn't have to convince the world. It is the sales manager or the New York state division manager who drives the Cadillac.

Because of its extreme status, it also holds magnetic attraction for people deprived in other areas of living. On the other hand, there are some of its former high status market who resent it for this specific attraction; it is now too common for them.

The other cars in the fine car field didn't have these overtones in any degree. Lincoln had a sporty aura, a luxury elegance. Chrysler Imperial was an expensive, modern car. But they were not the rainbow's end of the Horatio Alger dream like Cadillac.

Each car has its own distinct image which sets it off from the others. In addition to many other favorable attributes, Buick is seen as the car for successful people moving upward. It is an especially popular car among small-town lawyers, bankers and doctors who for

policy reasons wish to avoid any criticism for owning the highest priced cars.

The car purchase is comparable psychologically to building a house in the sense that it is a composite of many things. One element, such as color, may be balanced or combined with the model to produce a particular self-satisfying totality. Price, performance, styling, and any number of factors are offered and the buyer selects this, rejects that, establishing a composition of elements especially suitable to him.

There is no simple relationship, however, between kinds of buyers and kinds of cars. Any human is a complex of many motives—his practical aims, economic limitations, personal characteristics, his own particular social pressures. All of these motives should be seen as patterns with predominant and subsidiary meanings which may vary in countless combinations.

UNDERLYING ECONOMIC FACTORS

The basic set of appeals also existed fifteen years ago. It does not explain the enhanced desirability of the car, or the recognizable swing to brighter colors and sporty models, or why design has become the primary component in sales appeal. These elements are directly related to a visible change in our American style of life, to the underlying social and psychological currents which are causing such other drifts as the pronounced shift to informal living, the vast increase in hobbies and participation sports.

Since 1939, our tremendous economic gains have spurred family income far past the subsistence level. Historically, periods of widespread leisure and prosperity very often lead to: (1) a considerable increase in self-indulgent behavior; (2) a greater seeking for new thrills and stimulation; and (3) a desire for much more self-expression.

Whereas in other such periods of history the favored ones indulged themselves in art, magnificent estates with fountains and formal gardens, or stables of fine horses, today the automobile is the instrument best fitted for our whole society to give form to these motives. Where twenty years ago your conservative neighbor would never have dared to drive a lavender-flamingo ranch wagon, today the drifts in our living styles have made it conventional to use vivid colors as an outlet for self-expression.

This does not mean any increase in the number of attention-getters in the population or any decrease in the number of conservatives. It simply indicates that more people in all groups and types now seek

opportunity for self-expression. Besides such outlets as do-it-yourself activities around the house, new participation sports such as skiing, sailing, and pheasant-shooting, there are many elements about the automobile which permit the greatest number of people the greatest opportunity for self-expression, such as color, other-than-standard models, plush interiors, and extra gadgets.

In the 1930's, emphasis was on the depression and subsistence values of price, economy, and dependability. Today we look for products and services which embody these other motive areas of more individuation, exciting new stimulation, more glamor and color. The functional factors of comfort and sturdiness are far less appealing than designs which permit people to attribute these new values about themselves. Hence, the competitive battles are won, not by gear shifts and more head-room, but rather by exciting designs and new color combinations.

The automobile elbows other products aside because, more than anything else, it is the most self-expressive thing we have in this era when a premium is placed on self-expression, being acceptably different, being a colorful person within a framework of conformity.

IMPLICATIONS FOR ADVERTISING

In light of these study findings, it seems obvious that the most effective automobile advertising should utilize two basic sets of motivations, which in a different way could apply to any product: (1) the basic wish to ownership of the product, and (2) the personality of the particular brand. These are two different directions, but they both offer fertile fields for creative minds to harness. As a matter of fact, both levels of appeal can be in the same advertisement.

Much car advertising falls into the "name-and-a-claim-and-a-price" category, taking no recognition of the area of motives. Whatever impact this advertising has is purely utilizing the attraction of the car's present character. It concentrates on giving the car certain mechanical or stylistic features which people are supposed to value, with no regard for the pressures which caused him to want to spend several thousand dollars.

These are typical of the practical and expressive motivations which exist as potential copy areas in this particular field:

—The wish to be Somebody, to convey to others that we individually have interesting, colorful, effective personalities.

—To acquire a sense of financial power and importance. The auto is one of the more recognizable signs we have of financial success—the most important goal in the life of the American male.

—To accomplish transportation. Unlike many products such as beer or cigarettes, the automobile does have tangible practical purposes.

—To convey a sense of strength and capability, a feeling of personal power, a command over the environment.

—To provide socially acceptable outlets for basic emotional forces, at the same time symbolizing self-control.

—To show that one is adult and entitled to the privileges of adulthood. Not only does this seem important to teen-agers, but to various grown-ups who always feel they must prove their right to be considered adults, like short people.

—To extend one's life boundaries—having freedom of action at least subjectively, being independent in one's mind.

—To show social standing—the group or the rung on the ladder to which one belongs, or secretly hopes to belong.

—To show some personal superiority. Competitive rivalry underlies so many of our actions, trying subtly to show how we excel. Car ownership permits us in a variety of ways to impress our friends and work-associates. "My car has a better performance record; the model I'm driving is absolutely exciting; the new push-button gear-shift I have is wonderful and you don't have it."

—To express one's own individual personality and style of living.

—To celebrate the important personal changes in our lives. The car is a concrete means of making others aware that something significant has happened: getting married, an important promotion, a change in our circumstances from nose-on-the-grindstone to easy comfort, achieving maturity, and so forth.

Turning to the other facet of the advertising problem—selling any particular make—advertising has to reckon with the broad attitudes that exist toward the car, comprising its product personality. Advertising is most effective when presented in terms of this product personality, and when it is expanded to become more desirable to more different kinds of people.

The dimensions of these objectives can be stated this way: What is the current character of the brand in the minds of the general public, and what does the advertising do to build, modify or contradict this character? What is the long-run character the advertiser wants to establish, and how does the campaign contribute to this desired character?

It is highly important to stress—and this applies to any product or institution—that advertising must come close to fitting the existent product personality. Of course this can be widened, but for a well-known product, the process has to be done gradually in acceptable steps. A complete contradiction of the public image of any product is merely rejected.

People believe the advertising that fits their preconceptions. Again and again in this study, the persons interviewed about certain car advertising would sound off their opinions of the car first and halo

these likes and dislikes into their judgment of the advertising. A certain car that was considered rather stodgy, unexciting, for the Casper Milquetoasts of this world, was portrayed in one advertisement arriving at Sun Valley. It was a ludicrous situation to the people interviewed, asking them to accept that the mousey, ultraserved buyers of this car would ever venture to a sophisticated, gay resort.

In the broadest sense, the objective is to endow the car with the widest possible range of appeal. Unlike some products which are content with a thin segment of the population, the automobile aspires to mass appeal. Those cars with narrow personalities, either because of extreme features or single-track advertising, definitely limit themselves. This does not mean that each ad should be loaded with a multiplicity of appeals, but rather that by using different campaigns with different themes, entirely different audiences may view this particular product as most expressive of themselves.

Ford retains from its tin-lizzie past a recognizable market which values it as the most economical, the best for rugged wear, a plain farm car. On the other extreme, it is intensely attractive to youngsters and sophisticates for its flair, its power, its speed. It is very popular with high-income families wanting a second car. The choice it offers in ranch wagons and convertibles and station wagons has particular attraction for the smart set. Its range of color combinations constitutes an appeal for another type of individual. Each of these highly distinct personality groups sees some facet of the Ford personality as reflecting itself.

As biological organisms, all men are similar. But multiple pressures operate to differentiate us so that we look at the world through different eyes, with different values and beliefs, and different tastes. In the final analysis, advertising is the one means the manufacturer possesses which he can consciously control to represent his car in many ways to different people. The car is only a mechanical object with certain mechanical features for a price. Its success or failure in this day of technological near-perfection will result from its subjective desirabilities which are essentially unrelated to mechanics.

6. MOTIVATION RESEARCH—OPPORTUNITY OR DILEMMA?*

Here is a powerful indictment of the validity of psychological methods as applied to marketing research. In many ways it is the antithesis of the Dichter article: note the emphasis on the need for quantification, and the criticism of depth interview methods and projective techniques. At the same time, there is agreement even between these two extremes that psychological techniques can be useful. The disagreement lies in the scope of this use, Politz restricting it to the development of hypotheses while Dichter maintains that such techniques are sufficient in themselves.

Somewhat apart from the discussion of the validity of psychological techniques in commercial research is Politz's treatment of "true" answers and useful answers, a provocative subject in itself.

Because of the great variety of activities which are included under the term "motivation research," it is most difficult to evaluate objectively and thereby express a critical point of view of "motivation research" as a system. Nonetheless, advertisers, as well as members of the popular press, have their definite impressions as to what "motivation research" stands for and does. Since, to a large extent, these impressions are not compatible with sound research principles, this article will attempt to evaluate critically those which seem to be the more important. It is hoped that in doing so workers within, as well as those outside, the realm of "motivation research" will find it easier to offer aid to business and industry management without the subordination of either research or ethical principles.

MULTIPLICITY OF MOTIVATION

Among other things, "motivation research" has impressed people as overcoming the need for numbers. The term "nose counting" has been used to devalue the use of numbers. While this phrase may have a promotional appeal to lay audiences, it also violates scientific principle. Before doing away with counting, we had best ask whether an advertiser might be interested in the frequency with which a certain motivational phenomenon exists among consumers.

* Adapted from a talk by Alfred Politz, "Motivation Research—An Opportunity or Dilemma?" given at the Marquette Conference, 1955; later printed as "'Motivation Research' from a Research Viewpoint," *Public Opinion Quarterly*, Vol. 20 (Winter, 1956-57), pp. 663-73.

It is true that consumer studies consisting merely of statistical exercises have been labeled and passed off as scientific research. The employment of statistics in the absence of imaginative hypotheses often leaves a manufacturer with the feeling of "so what?" (In this context, reference is not intended to the market statistics provided by organizations which concern themselves with specialized services of a statistical nature.) Too often the advertiser has contracted for action-guiding consumer research and in turn has received a considerable number of statistical tables irrelevant to a practical solution of marketing or advertising problems. The embracement of the rather confused activities subsumed by the term "motivation research" may have been an emotional reaction to the unproductiveness of such an approach. It became sophisticated to ridicule "nose counting" and to concentrate on the seeking out of "the real motive" underlying consumer behavior. It was assumed that if we could identify "the real motive," that is, "the reason why" product X is bought, then we could capitalize upon this information to increase product X sales. We must stop to consider what is meant by "the reason why."

If it were possible to reduce the causal relationships accounting for consumer behavior to a single, easily describable reason, matters would indeed be very simple. The linguistic habit of using the singular form, "reason why," may well serve to reinforce the prevalent expectation of a single or at least a single major reason accounting for any phenomenon. This expectation is so entrenched that advertisers have been led to believe in the existence of a "formula" which can identify the singular "reason why" for specific instances of consumer behavior. A researcher with even minimum knowledge of the cause-effect relationships could not begin to suspect the existence of this panacea. Rather, he would expect that the mere belief in such a panacea will obscure the real situation. Any one specific piece of consumer behavior goes back to a multitude of psychological and mechanical causes. It is the specific combination and the relative strength of the various causes which lead to a specific purchase. To take Coca-Cola as an example, everyone will agree that the taste of the drink contributes to the purchase. Taste, thereby, is a reason for buying and drinking, a reason of such magnitude that a change of the taste may destroy the whole brand. There is, however, the taste of Coca-Cola's competitors. The taste of a competitive drink may pull consumers away from Coca-Cola or drive them to Coca-Cola. What is the real reason for drinking or not drinking Coca-Cola?

It is obvious that the price of Coca-Cola and the price of competitive drinks both constitute new reasons. There is the fact that some people are strongly motivated to drink Coke because they are thirsty. There is the social symbolism of a Coke among teenagers. There is the relative strength of the attraction Coca-Cola offers as a social catalyst when boy meets girl, the convenience of obtaining Coca-Cola, the convenience of obtaining other drinks, the convenience of carrying, storing, and serving the liquid from different bottle sizes. There is the feeling of a "lift" connected with drinking a Coke. There is a feeling of relaxation, of rest, connected with drinking a Coke. Each of these causes is sufficient in strength to raise or lower the frequency of purchases. It is the privilege of the nonresearcher to talk loosely about the real reason or the reason why. It is the obligation of the researcher to recognize that a multitude of causes (some of them may be called reasons) lead to the effect: the consumption of Coca-Cola. If we intend to discover the relative strength of the contributing causes, we have to measure them. If we want to measure, we must use numbers; we must denote magnitudes. We must quantify as the researchers in the physical sciences must quantify.

The fallacy which underlies the talk about "the reason why" people buy has done considerable damage to productive consumer research, and thereby has done much damage to success in marketing. Causes act in different directions, partially benefiting and partially retarding sales. If we do not know the relative strength of these causes, we cannot compute the effect of a practical marketing action. One can perform a useful function in marketing by getting a good idea. But not everything that is good must be called "research."

The causes behind consumers' actions are infinite. Some causes, as we have seen, are a matter of physical circumstances, while others are of a subjective nature. Cutting across this classification of physical and subjective causes is a classification of controllable causes. The problem of the researcher is not to find causes, but to find controllable causes. Controllable causes are those upon which we are able to take action. This knowledge of controllable causes, however, must be part and parcel of an over-all view of the manufacturer's problem. The adequacy of the over-all view is decisive in giving validity to the design of the research.

The purpose of consumer research is to tell the manufacturer or advertiser what to do in order to sell more of his product or services to the general public. It includes not only research designed to investigate the motives of consumers, but also research designed to

investigate such other psychological aspects as attitudes and, particularly, ability to learn. It includes also a study of such mechanical and circumstantial factors affecting the sales of a brand or product as availability, distribution, and relative price of the product or brand. It recognizes that behind every consumer's act of buying or not buying there are a multiplicity of causes and not a unique cause. It is the function of this type of research to discover those causes which can be economically controlled by the advertiser in order to increase his sales.

Thus, if consumer research is to be useful rather than entertaining, it must lead to predictions. The predictions involve risk—including the risk that the possible inadequacy of the research work will eventually be detected. Some think this risk can be avoided if research confines itself only to measurement of present situations or, as this is often called, the collection of facts. This concept, however, is both selfish and illusionary. It does not change the principal issue. It only suggests a division of labor in which the professional researcher is careful not to participate in drawing productive conclusions from the statistical data he has collected. The conclusions and, therefore, the predictions are left for someone else to make—the sales executive, the manufacturer, the advertising director, or management in general. Guiding the marketer means providing a basis for making decisions. A decision entails the expectation of a specific result. Therefore, a decision entails a prediction. Worthwhile marketing research arrives at the causal conclusion that action A will lead to result B. A decision can then be made to take action A with the knowledge that result B will occur.

DEFINITIONS OF "MOTIVATION RESEARCH"

Consumer research is a multiphase process. Following the explicit statement of the problem during discussions with the client, a search for ideas or hunches relevant to the problem begins. These ideas or hunches then receive rigorous expression in terms of hypotheses to be put to test. The procedures to be followed in testing these hypotheses must then be carefully spelled out, considerable attention being given to the anticipated analysis of the data to be accumulated and the relationship required between sampling operations and questionnaire construction so that the analysis may take place. The field survey is then undertaken and the resultant data tabulated. The data are analyzed and interpreted, conclusions drawn, and the client is presented with a report embodying the analyzed

data with specific recommendations for action. Quite superficially, this takes one through the process of consumer research.

It is at the stage of searching for ideas or hunches that one makes use of the activities subsumed under the term "motivation research." These activities may or may not involve the use of projective techniques and the semidirective or "depth" interview. It should be noted that even in this search for ideas and hunches, direct questioning procedures and statistical techniques are also employed. That is, even in the search for ideas and hunches, *all* of what is popularly known as "motivation research" comprises but a part of the operation.

What should the term "motivation research" correctly signify? It should designate research aimed at a better understanding of motivation. This consideration would be parallel to an interpretation of chemical research as research aimed at a better understanding of chemical reactions. In the case of "motivation research," however, a freakish situation has developed in which a few particular tools, which might or might not work, have assumed the name of the job. To give the name "motivation research" to specific instruments, which in the best of cases contribute to consumer or marketing research, is comparable to a situation where one adds a new instrument to a symphony orchestra and names this instrument "music."

If the use of incomplete sentences, for example, is called "motivation research," then, by the rules of the English language, the use of direct questions must be called "motivation research." Therefore, either the term "motivation research" should be dropped altogether and the techniques which erroneously have been called "motivation research" should be referred to simply as projective techniques, or all research aimed at an understanding of consumers' motives should be referred to as "motivation research" regardless of the techniques employed. In the latter case, about half or three-fourths of advertising and marketing research must be called "motivation research." In such a situation, the term "motivation research" will take on a clear meaning. Advertisers and journalists will be less confused and less excited. They will become acquainted with the fact that in an attempt to understand consumers one may use a variety of techniques. The news value of the term "motivation research," of course, will be substantially reduced if the advertiser and the press become aware of the fact that "motivation research" is but a very small fraction of real consumer research. If the term "motivation research" is applied to research aimed at consumer motivation, then many advertising agencies will be relieved of the pressure of a misnomer. They need

not go into the motion of putting ink blots before prospective consumers just to prove to an unscientific and confused advertiser that they are interested in consumer motives.

The title "motivation research" consists of two terms: "motivation" and "research." The function of research, in this instance, is to help the manufacturer sell more products to more people. This usually means dealing with thousands and millions of people whose motivation is of interest to us. Advertising must concern itself with those few psychological conditions which millions have in common so that the same message, that is, the same ad, the same commercial, can be directed to millions of different people. Under these conditions, one should assume that research related to consumers must concern itself with millions. On the other hand, the concern we have with millions does not necessarily lead to the need for interviewing millions. All we need do is interview enough people to project findings with satisfactory reliability to the millions.

If the number of respondents interviewed is either too small or unrepresentative of the whole, the investigation is not necessarily done in vain. The investigation might have value as a thought stimulator for creative men. It may stimulate the imagination of copywriters or sales managers. Ideas which result from this stimulation may lead to successful marketing activities. We must be aware, however, that cases of this type are not research. Rather, such investigations should be considered as motivation conjecturing or motivation guessing or hunch hunting instead of "motivation research."

We must not underestimate the importance of such investigations in producing hypotheses. No genuine research can hope for productive success unless it starts somewhere with hypotheses—that is, rigorously formulated hunches. In conducting research, the test of these hypotheses must be based on an investigation which is designed in such a manner that our instrument is not only able to confirm the hypotheses but equally capable of rejecting them. It is most erroneous to take the view that research, in order to be objective, must not have opinions or hunches. The opposite is true, the more opinions and the more ideas, the greater the usefulness of the research. There is no limitation to the importance of hunches provided they are treated as hypotheses and not as findings. A very large portion of action-provoking findings which we submit to manufacturers are identical to the hunches with which we start out. We can make recommendations for specific action, however, only after testing the hypotheses by objective procedures, and we can consider them findings only

after application of the test. Without research we have ideas and opinions, some of which are right, some wrong. The purpose of research is to test these ideas and determine by measurement which are right and which are wrong. If the use of an adequate and representative sample shows that the motivational situation guessed from a few informal or so-called depth interviews exists in only 3 per cent of the consumers, it might be uneconomical for the advertiser to act upon it. If, on the other hand, the conjectured motivational cause exists with 90 per cent of the consumers, there will be no doubt that the advertiser should act upon it.

Use of the term "consumer research" rather than "motivational research" at least permits the implication that psychological causes, other than motives, and also mechanical causes, will be considered in any investigation. Only a part of human behavior can be explained by subjective causes. The behavior of a sailor who is washed overboard by a high wave is not efficiently described if one omits the roll of the wave from the causal consideration. The consumption of Esso gasoline by someone who parks in the Rockefeller Center garage, ignorant of the fact that this garage has an Esso pump to fill the tanks of the monthly parkers, can be sufficiently understood only if mechanical causes are taken into consideration.

ARE TRUE ANSWERS ALWAYS USEFUL ANSWERS?

The attempt to transfer clinical techniques to consumer interviewing for the purpose of developing marketing ideas is justifiable. Clinical work, as well as advanced consumer research, has demonstrated that an understanding of a person's response does not necessarily depend on the logical content of his statement. The problem here is not to get a true answer or an unbiased answer. Rather, the problem is to observe responses to definable, controlled stimuli and then to draw conclusions from the variety of responses. The burden of solution of the problem rests on the experimenter. The attempt to introduce clinical techniques into consumer interviewing may help to dramatize the fact that the logic of a response which one obtains is relatively unimportant.

The usefulness of an answer is in many cases independent of the truth of the answer. This constitutes the difference between research and daily life situations where one expects, when questioning a friend, to obtain true answers. It is a disregard of research principles to carry the function of a question in daily life into motivational investigations. It is the privilege and obligation of objective research to use any means (within ethical limits) to obtain a variety of responses

which, when treated as a system of premises, lead to a system of conclusions. All questions treated in this manner must be understood as controlled stimuli rather than as a device to produce the truth. The elementary prejudice that responses in themselves have to be "true" before they can be used in research activity has frequently handicapped opinion and attitude research, and, it appears, recently has led so-called "motivation research" back into a twenty-year-old trap.

It cannot be surprising that in predictive research the operational rules are similar to those in the physical sciences where the relationships between experiment and prediction have been understood for more than three centuries. Answers to questions can no longer be considered as solutions to problems, but rather responses to specific stimuli. Definition and control of the stimulus determines the interpretation of the response. Seen in this light, the question of whether the response is biased or unbiased is no longer relevant. Responses are reactions observed under specifiable conditions. The conditions are the environment at the time of the interview, the appearance of the interviewer, the "internal state" of the subject, and the questions asked. If a biased question is needed to get a useful answer and if a biased answer is needed to serve as a basis of prediction, we are obtaining valid research.

There are situations, of course, in which the truth of the answer is desirable. But we increase our opportunities to understand consumers if we also put untrue answers to use. For example, at one time the question, "What brand of refrigerator will you buy next?" produced more votes for General Electric than for Frigidaire, but Frigidaire sold more than General Electric. This is a case where respondents told the truth. The respondent genuinely thought that the next refrigerator he purchased would be a General Electric. The stimuli, as they existed at the time of the interview, elicited opinions that General Electric was the best refrigerator. On the other hand, the stimuli which came into play when preparing for the actual purchase were a different matter. The fact that General Electric was considered best did not conflict with a curiosity to see other refrigerators before purchasing. Window shopping became pleasant. Reading the copies of ads related to refrigerators took on interest. The respondent learned about features of Frigidaires, Kelvinators, and other refrigerators; he listened to salesmen's arguments. The totality of stimuli which were altogether different from the interview stimuli created an opinion more favorable to Frigidaire than to General Electric. This shows

that only by recognizing the limitations of the truth can we avoid being misled. It is a misunderstanding of the motivational mechanism if we believed that market problems can be solved by obtaining from the consumer his true attitude.

Being misled by the answer given to questions has erroneously contributed to the belief that the untruth of the simple answer to a direct question is only the superficial product of a superficial interviewing procedure. The hope developed that by digging "deeper" into the consumer's mind one would finally find the truth under the layer of initial superficial responses. The popularity of such an illusion is increased by the connotations that go with the word "truth." To get at the truth is, of course, a moral and practical requirement in daily life situations. Part of the respect for "truth" is then carried over into specific situations where the respectable meaning of this word is irrelevant. From a person's lie one can arrive at the true finding that the person is a liar. In all research dealing with people, the truth has to be calculated, has to be concluded, has to be discovered by the researcher. But the truth doesn't have to be expressed by respondents subjected to the research. The old hope of getting the true opinion of a person by unbiased questioning is the twin brother of the new hope of getting at the true reason by avoiding direct questions.

THE "DEPTH" OF "DEPTH INTERVIEWS"

The semidirective or "depth" interview is frequently considered a projective technique but, in reality, is not. It may utilize projective techniques to glean information from the respondents, but in and of itself it is nothing more than an informal, semidirective interview procedure. Use of the word "depth" implies that the unconscious is tapped and hidden motives identified. The unconscious postulated by Sigmund Freud, however, could not possibly be the unconscious alluded to by "motivation research." If one sees any similarity between semidirective interviews and a psychoanalytic session, he is confining himself to a very superficial consideration. In both situations we find two people who talk to each other, one of them somewhat helped along by the other. The real problem is what goes on in a psychological sense. Clarification of this point requires a better understanding of the motivational background that leads to the conversation between two people in consumer research and two people in a clinical situation.

Freud strongly emphasized that a psychoanalyst should not accept patients who come only because of the strong pressure put on them

by their relatives and friends. Psychoanalytic treatment rests upon the assumption that the neurotic became a neurotic because he did not want to accept the recognition of painful reality within the conscious part of his mind. The aim of psychoanalytic treatment, therefore, consists of a re-admission of the repressed element into the conscious part of the mind. Of course, this re-admission is painful. Weeks and months of conversation with the patient gradually enable the analyst, with the help of symbolic interpretations, to identify the repressed elements. These meetings finally enable the analyst to help the patient accept the pain of recognizing reality as a lesser evil than the pain of neurosis. As long as the pain resulting from the neurotic ailment is not strong enough to drive the sufferer voluntarily into the office of the psychoanalyst, it cannot be hoped that the neurotic will accept the pain of the psychoanalytic treatment; that is, he will not accept the invitation to open the passage to his "unconscious" in the Freudian sense.

In consumer interviews the motivational mechanism is the opposite. The interviewee does not look for help. It is the researcher, the interviewer, who more or less imposes himself upon the interviewee. It is the researcher who goes to the consumer for a selfish motive—that is, to get money, power, or fame. We do not try to help the consumer whom we subject to our tests. We might be thoroughly convinced that in the long run our activities lead to the improvement of every consumer's happiness. The general happiness increase which may result from better living through our research is very small, however, for any one individual whom we happen to select for our tests. If an individual is willing to comply, he does so on the basis of less powerful motives than relief from pain. Consequently, it cannot be hoped that under these motivational circumstances the "depth of the unconscious" can be uncovered.

Freud conceived of repressed and suppressed forces so powerful that they distorted the personality. Can these be the forces tapped by the "depth" interviewer in a consumer interview? Unlikely. Freud's unconscious took months of probing before the patient could verbalize its content. In the few hours of conversation to which the consumer is subjected, there couldn't possibly be a tap of this repressed material. Consequently, the interview has no depth other than the "depth" that any conversation with friends, journalists, lawyers can have. The responses elicited either directly or indirectly are, in the accepted sense of the word, conscious responses. If one insists on the use of the word "depth," then it would be more correct to speak of depth

interpretation rather than depth interview. As a matter of fact, leading psychoanalysts, loathing sensationalism even in their own clinical realm, insist on the very sincere and correct term "depth interpretation."¹ This term denotes that the crucial factor is the clinician's interpretation rather than an individual's responses! This seemingly minor terminological point is of consequence.

The trap that many a psychological discussion falls into is a misunderstanding of the unconscious as the hiding place of things of which a person is not conscious. Needless to say, Freud's analytical powers kept him aloof of such fallacies. A person is not conscious of the light intensity diminishing by the square of the distance between the eye and the source of light. Not being conscious of it does not mean that there is a picture of this phenomenon in his unconscious. A nonexperimenter is not conscious of the causes which lead him to incorrectly estimate sizes in well-known optical illusions. Yet these causes are not part of his unconscious. The clinician is obligated to deal with the phenomena located in the unconscious of the Freudian sense. The consumer researcher cannot confine himself to the consideration of motivational phenomena, conscious or unconscious. The consumer researcher's task is easier than that of the clinician in that he does not try to rescue people from painful conflicts. The consumer researcher's task is more difficult in being forced to predict and control consumer behavior even though the causes of such behavior are deeper than the "depths of the unconscious."

There are situations where the nature of the problem is such that utilization of informal, or rather, semidirective interviews are advisable. For example, in investigating problems related to products such as sanitary napkins, the use of semidirective interviews tends to alleviate the embarrassment, inhibitions, and taboos surrounding the discussion of product usage. Too often, however, temptation to use informal interviews exceeds the justification of their use. Underlying this state of affairs is the problem of experimental design. If a researcher lacks the interest or ability to design a questionnaire, he may be too willing to resort to semidirective interviews. Insufficient understanding of the problem of experimental design to a large extent seems to be responsible for the idea that a question is less apt to "dig into the depth" of a consumer's mind than the allegedly free associations of a semidirective interview. While there are situations where the semidirective interview is preferable, it must be emphasized that in other situations a rigidly

¹ E. Glover, *The Technique of Psycho-Analysis* (New York: International Universities Press, Inc., 1955).

structured interview based on a rigid design is capable of uncovering a causal mechanism so deeply "buried" that no "depth" interview could reveal it.

HOW VALID ARE PROJECTIVE TECHNIQUES?

Of the activities subsumed under the term "motivation research," those activities dealing with the use of projective techniques appear to cause the greatest amount of excitement among members of the advertising and journalistic professions. These projective techniques are techniques which were designed for clinical purposes. Such techniques were presumably introduced into the field of consumer research as supplements to the tools available to the consumer researcher. There are some indications, however, that these clinical or projective techniques were introduced not as supplementary tools, but rather as the replacement of one type of instrument by another—that is, the use of sentence completion, cartoons, and so forth, in lieu of direct questions.

The fact that projective techniques and psychological terminology with constant references to clinical situations are used in so-called "motivation research" has given "motivation research" an aura of scientific glamor. Psychologists with scientific knowledge, however, are well aware of the low validity of these techniques. Many experiments have been conducted and have been documented in the scientific literature which indicate that these techniques have severe limitations even for diagnostic purposes. Professor Robert J. Williams² and Professor Joseph Zubin,³ both of Columbia University, are but two of a number of researchers in this area who have commented and reported on the lack of scientific validity in projective techniques. Two investigators, Holtzman and Sells, of the USAF School of Aviation Medicine, designed an experiment to determine whether a group of clinical psychologists could predict successful completion of a flight training course on the basis of a battery of projective tests. They conclude that: "The clinical assessments of beginning aviation cadets have no relationship to a criterion of adjustment in the basic flight-training program."⁴

² Professor Williams commented on the limitations of projective techniques in a speech delivered on April 24, 1956 before the Metropolitan New York Association of Applied Psychology.

³ J. Zubin, "Failures of the Rorschach Technique," *Journal of the Projective Techniques*, Vol. 18 (September, 1954), pp. 303-15.

⁴ W. H. Holtzman and S. B. Sells, "Prediction of Flying Success by Clinical Analysis of Test Protocols," *Journal of Abnormal and Social Psychology*, Vol. 49 (October, 1954), pp. 485-90.

Because the clinician aims at findings related to a single individual, it is possible that a battery of tests of low reliability and low validity may yield some clue and an improved guess. In the long contact over weeks and months between the clinician and patient, a fairly good picture might accumulate from the mixture of projective techniques, direct questioning, and silent observations. Nonetheless, Jean Walker Macfarlane and Read D. Tubdenham, both professors of psychology at the University of California, have the following to say on the validity of the projective techniques:

In this volume, whose very size and scope reflect the vitality of projective techniques, the present chapter may seem to offer remarkably little solid evidence for the "scientific" validity of these devices, and to present an embarrassing number of unsolved problems in establishing their scientific worth. A clinician using projective techniques soon develops a strong conviction that they, or at least the ones in which he has had rich experience, present important data about persons and personality dynamics. Yet, many workers with such strong convictions, who also have the equipment of disciplined scientific methods at their disposal, have, to their dismay, found little in their own research findings or in those of others to justify their pre-research enthusiasm.⁵

Sigmund Freud's genius developed the most amazing instrument for the rediscovery of repressed emotions within a single person. Freud was dealing with psychological forces of great power. The consumer researcher has to deal largely with forces (including motives) of low intensity. For example, a motorist approaching a Shell station may have in mind to turn in and buy gas, but he notices that another car is being serviced. He therefore decides to pass the station and turn in at the next one, which happens to sell Gulf gasoline. Whatever minute reason motivates him to buy Shell or Gulf is typical of the problems in real marketing. It is the multitude of such small forces multiplied by the millions of consumers which makes one brand a success and another a failure. Therefore, refined techniques had to be developed in consumer research. These techniques utilize statistical methods and the principles of experimental design.

An analogous situation would be the case where one had to determine the thickness of paper coming out of a paper machine, and had available only a ruler calibrated in inches. This is a very crude instrument to be used directly for the task at hand. Making use of experimental design, one would stack a pile of, say, 1,000 sheets of paper.

⁵ H. H. Anderson and G. L. Anderson (eds.), *An Introduction to Projective Techniques* (New York: Prentice-Hall, 1951), Chap. 2.

He would measure the height of this pile and divide this measure by 1,000, thus determining the thickness of a sheet. Anyone who cannot see the value of design and numbers will be deprived of ever being able to discover and measure motivational forces of intensities so low that Freudian techniques are bound to fail, just as they will be deprived of measuring the width of thin paper. For the discovery and appraisal of forces that motivate the consumer to consume the one or the other brand, the Freudian tools are as crude and inadequate as a mallet applied to a wristwatch.

The advanced researcher interested in consumer motivation makes use of the enormous magnifying power that goes with the statistical instrument. Of course, Freud and his school could not have been expected to develop these tools. They had no need for them. An attempt to cure a neurotic starts with a consideration of the one individual's characteristics; his specific conflicts may be different from those of all other individuals. On the other hand, in consumer research we are concerned with those motivational aspects which are related to thousands or millions of people. Our whole economic system rests on the fact that the *same* product in the *same* package is to be sold to millions of different people and that the *same* television commercial will be listened to by millions of different people. Therefore, we do not achieve anything if we discover specific reasons for specific individuals. We must discover reasons which a sufficient number of consumers *have in common* so that a manufacturer can sell a standardized product to them and so that an advertiser can afford to address them with the same copy.

SUMMARY

We may say that a consumer researcher has to pay a great deal of attention to forces of low intensity. A clinician, on the other hand, is dealing with forces so powerful that they create mental illness in a person. A consumer researcher has to discover forces which so many consumers have in common that the marketer can, with one action, appeal to the millions. On the other hand, the clinician must interest himself with the set of forces which might make his patient different from all other people in the world. The low intensity of the forces operating in the consumer field makes it necessary to develop refined statistical instruments before the forces can be discovered. On the other hand, the fact that clinical therapy is applied to an individual makes the application of such a statistical instrument unnecessary.

"Motivation research," or better, the employment of projective tech-

niques, can play a preresearch role in marketing or consumer research. With or without the activities subsumed under the term "motivation research," one can develop ideas and hunches. Without ideas there could be no research, for in research we must have ideas (hypotheses) to put to test. It must be stressed, however, that "motivation research," as presented to advertisers, is hypotheses hunting, not hypotheses testing; it is preresearch, not research. The dramatic proclamations that accompanied its arrival on the marketing scene stimulated considerable interest in "motivation research." If an equivalent amount of interest could be stimulated in the role of experimental design in consumer research, advertisers would benefit from the guidance of genuine research, and from proof instead of possibly unresearched opinions.

7. RATIONALITY AND IRRATIONALITY IN MOTIVATION RESEARCH*

This indictment of psychological techniques in marketing research is, if anything, more direct than the preceding one. Many of the points are, in fact, similar, but the approach is from more of a practical operating level, and it will be seen that the author is forthright in his criticisms.

It is interesting to note the attempt to distinguish between "motivation research" and "motivation analysis," ascribing to the latter term the meaning that many commercial psychologists have tried to impart to "motivation research."

The purpose of this paper is threefold: (1) to describe what is meant by the term "motivation research," (2) to explain the rational and irrational aspects of motivation analysis as it is practiced today, and (3) to provide some current examples so that any businessman will be able to recognize with which school he is dealing, if he is approached by any person or firm purporting to do motivation research, and not find himself in the predicament of so many companies today, who, having paid a lot of money for a motivation research study, are now more confused than they were before.

Typical of the confusion surrounding motivation research is the fact

* Adapted from a talk, "Rationality and Irrationality in Motivation Research," by L. Edward Scriven (based on ideas developed by Fred T. Scheier, Research Director), A. J. Wood and Company, given to the Washington, D.C. Chapter of the American Marketing Association, June 23, 1955.

that the practitioners themselves have difficulty understanding each other, because of the lack of standard terminology. The semantics problem has been such that the Advertising Research Foundation felt it necessary to publish a glossary of terms for the benefit of its members. Typical of this same confusion is the fact that the name "motivation research" is a misnomer for what should be called "motivation analysis."

Motivation analysis can be defined simply as a set of tools, borrowed from the fields of psychology and sociology, to uncover and evaluate the motives or drives that are back of the human behavior of the consumer market. Once the nose-counting job has been completed—that is, the determination of how many people do what—the next problem is to determine *why*. Motivation analysis is the means by which the market researcher attempts to do this. How successful he is depends, as shall be seen, on whether he proceeds rationally with his new-found psychological tool, or whether he borrows from the irrational methodology and ends up with a completely confusing batch of data.

Every marketing research man today must pretend he knows all about motivation analysis, believes in it, and is in favor of it, whether he is or not. Probably most research men, and most marketing executives, if they understood that there are two distinct types of motivation analysis—the rational and the irrational—would insist on knowing which was to be discussed before committing themselves as to their attitudes. But they don't. They think all motivation analysis is alike—maybe good, maybe bad—but they must pretend they are in favor of it for fear of being considered old-fashioned and ignorant of modern methods. One market research man, quoted recently in *Time* magazine, said, "It's like asking if you are against sin. You can't afford to say no."

Some marketing executives feel that the whole business of motivation analysis is simply a hodge-podge of jabberwocky, or the line of a glib psycho-salesman bent on selling fifty "depth" interviews for \$50,000, because quite a few manufacturers have been taken in by just such a line—ending up poorer, but not wiser—because they didn't know how to separate the rational from the irrational in advance.

The psychological approach to marketing research does not offer a panacea for the marketing problems we have—even when rational methods are used. It does provide a starting point from which the statistical researcher may determine the relative importance of various motives and attitudes. Such knowledge can be of tremendous importance to the marketing executive.

If we can accept the above explanation as a definition of motivation analysis, let us look at the second problem, the explanation of the rational and irrational in motivation analysis as it exists today.

The introduction of psychological factors to the field of marketing research is good. All marketing executives definitely need to know the attitudes and other factors which influence people to buy. If we can determine these influential factors and they are the sort of thing we can do something about, then obviously we can do something to increase the number of people exposed to the right influence to increase our sales. The question is how can we find out what psychological factors motivate masses of people.

The clinical psychologist long since found that every individual has his own complex personality which controls his motivations. As has been pointed out, personality operates on at least three levels. There is the outer level with which the world is familiar. Thus, we say a person is calm or excitable. There is the conscious inner level of personality which is the person's private world of secret plans, wishful thinking, and daydreaming. Then, there is the subconscious level where are located the strong forces which the individual himself may not recognize. The clinical psychologist attempts to penetrate into these second and third levels to reveal true motivations.

The clinical psychologist generally uses three methods in his probing procedure: (1) depth interviews, (2) focused group interviews, and (3) projective techniques. In almost every case, the clinical psychologist, when he attempts to apply his methods to determine motivations in the marketing field, carries over three principles of methodology by which his work is identified. The first is the conviction that it is only by the introduction of these clinical methods that the reasons back of someone's behavior can be discovered and recognized. The second is that these psychological factors can be uncovered only by one of the three listed techniques. Third, a small sample must be used. It would be entirely too difficult and certainly too costly to use a big (or adequate) sample. Statistical sampling methods are not required.

THE SINGLE-CASE APPROACH

The clinical psychologist, in his own field of endeavor, deals with a single case, his patient, and here he relies almost completely on the depth interview to expose the motivations of that patient. When he moves over into the area of mass motivations, he usually tries to use these same methods of depth interviews, augmenting them with focused group interviewing and with projective techniques. These latter

two methods, however, are so subject to questioning in marketing research that we shall not dwell on them in this paper, but rather discuss the utilization of depth interviewing which is employed by all clinical psychologists attempting to do market research work.

The depth interview represents the type of probing used by the clinical psychologists. It does have several advantages. It does give the researcher an opportunity to uncover new motivations. It does permit the development of a motivational pattern with respect to the selection of a given product or brand. It does provide stimulating new ideas to the researcher.

But it also has a lot of weaknesses and dangers. The first is that, even with a two- or three-hour interview with each individual, the researcher is attempting to do what the clinical psychologist may take weeks or months to accomplish. A second weakness is that most of the people who practice depth interviewing are not trained psychologists and therefore do not have a systematic approach in conducting the interview. An even greater weakness is that no quantifiable data are obtained in the depth interview procedure; therefore, when it comes to drawing conclusions from depth interviews, the conclusion is likely to be simply the judgment of the researcher. Thus, different conclusions will often be found by two different researchers in the same situation, as in the example to be shown later. There is no possibility of verification.

It might be added that many people claim to do depth interviewing who merely conduct long unstructured interviews. These are simply informal interviews which can be very useful in determining ideas for a new hypothesis, but they certainly are not true depth interviews.

Recognizing the weaknesses of depth interviewing as a means of getting at mass motivations, some psychologists have used focused group interviewing. Here a small number, maybe a dozen, presumably representative consumers are brought together for an informal group discussion of a given question. An attempt is made to use the same procedures as in a depth interview with this small focused group. It does have certain advantages, but it still does not provide information from a scientific sample, nor does it provide any possibility of quantification or the determination of the relative importance of any motives which may be uncovered.

Space will not be taken here to discuss other means by which the psychologist attempts to get at the underlying motives other than to mention such things as word associations, sentence completions, picture recognitions, and Rorschach techniques, to list a few. Suffice to say that

the clinical psychologist in his own work is faced with the necessity of prophesying behavior on the basis of phenomena observed in a single patient. For the single patient, this may be all right. The clinical psychologist may say that, due to the presence of certain observed phenomena, he can state definitely that this individual patient will violate his parole, fail in college, become a public charge, and so on.

To determine the importance of any motivating factor in the marketing field, however, we must be able to determine the frequency with which that motivation occurs among a true cross section of the buying public. Thus, we cannot rely on information from a depth interview with one individual or even with a group of individuals. We must have facts which have been obtained through standardized questioning procedures from subjects selected from a true probability sample, and we must be able to compare the frequency of one such fact with another so that the marketing executive can say that one is important and another is not.

We may conclude, therefore, that the attempt to apply the methods of the clinical psychologist to marketing research are irrational and doomed to failure.

RATIONAL VERSUS IRRATIONAL MOTIVES

The Freudian psychologist believes that all important motives are buried in the person's subconscious and related somehow to the sex urge. This type of psychologist might say that the round cake of soap is best, and that people will buy it because its roundness reminds them, subconsciously, of the roundness of a woman's breast and its smoothness to the touch of their hands. The rational reason will be that the square cake of soap is best because it doesn't slip out of the hand. A rational approach to the problem of the best cake of soap will lead to the use of a rational hypothesis, such as, "a square cake of soap is best because it won't slip out of the hand." By the use of proper statistical methods to determine the frequencies and rank orders of all the factors which may contribute to this hypothesis, the soap manufacturer can determine the answer to his problem. As a matter of fact, this has been done. Today, there are very few oval bars of soap on the market. It has been determined some time ago that the rational aspects of this problem are the meaningful ones. The proof has been in a long record of sales.

The rational analyst knows that the determination of mass motivations must depend not only upon a standardized questionnaire used among a scientifically accurate probability sample of the buying people,

but also standardized questionnaires which are themselves built around rational reasons.

The problem is even more complicated by the fact that it is only rarely that a person does something because of one motivation. There are usually multiple factors or impulses back of a given action. For instance, if we try to find out why a woman bought a dress, possibly she bought it because of style, or color, or design, or fabric, or price, or because she needed a new dress or, possibly, because of *all* of these reasons. It is only by establishing the frequency and rank order of each of these motivations that we can possibly know their relative importance.

The rational analyst has found that he doesn't need the depth interview; in fact, that he can't use it. He may in his informal and exploratory work, in planning a survey, use an unstructured or open-end question procedure to help him develop the ideas which become the basis for his research hypothesis. Most market research men and research agencies have been doing this for many years. Many of them have been thoroughly conversant with motivation analysis for many years, but they have also been well aware of the fact that it is only by the statistical quantification of research findings that useful information can be developed.

Thus, we may state that the basic distinction between rationality and irrationality in motivation analysis is the distinction between the clinical single-case approach to determine motives and the statistical approach. The statistical approach determines the frequency and objective rank order of such rational reasons as: tastes right, serves its purpose best, is the most effective ad, and so on. Some of these ideas may be crystallized by a look at one or two actual case histories. Two recent cases show what happens when the clinical psychologist is employed and when he applies the clinician's techniques in an attempt to answer a marketing problem.

The following case was reported by the *New York Post*, June 1, 1955: Dichter and Vicary were employed independently by two separate groups in the prune industry to find the causes for the sluggish market. Dichter found some startling reasons why people disliked prunes. He found that the prune was a "dried-out, worn-out symbol of old age;" the prune "fails to give security;" it's "a plebian food without prestige," and so forth. Dichter recommended, therefore, that prunes should be renamed "black diamonds"—"surround prunes with an aura of preciousness and desirability." Meanwhile, Vicary reported that Americans have an "emotional block" about prunes' "laxa-

tive connotations." And Vicary recommended a blunt solution, "Exploit the core of the market by advertising the 'laxative features' and don't pussy-foot about this angle, either."

Which should the prune people believe? They do not know and no one can tell them unless these ideas which the two motivational analysts have turned up can be statistically quantified among a proper probability sample.

Here is a second case: A well-known manufacturer submitted an identical set of problems, separately, to two of the leading commercial practitioners of "motivation analysis." The questions submitted were: 1. What is the present status of this corporate personality of ours? 2. What should the role be, *a.* as a service personality, and *b.* as a sales influence?

Practitioner *A* conducted 200 depth interviews in various parts of the country. *B* conducted about 200 in one Midwestern city. Both *A* and *B* applied projective techniques to the situation. *A* reported the corporate personality as "warm," "understanding," "real." *B* reported the corporate personality as "cool," "aloof," "abstract."

But this contradictory confusion was not all. In answer to the second question, *A* found that the corporation "cannot afford to be tied up too specifically" with the products, and warned that the personality would "fall down on the job as far as psychological role" is concerned if used as a "sales device." *B*, on the other hand, found a large proportion of customers and prospects "favorably disposed" toward the corporate personality "acting as a salesman." The findings in both reports appeared to be based on the personal opinions of the analyst rather than on the reactions, attitudes and beliefs of a dependable sample of customers and prospects.

Here again, the irrational approach led to nothing but confusion in the mind of the client. It would be quite possible, however, to take the findings of *A* and *B* and, through the use of a standardized questionnaire in a proper sample, determine the relative importance of each finding among a cross section of the public and advise the manufacturer that finding number 1 was relatively unimportant; that number 2 was typical of 85 per cent of the respondents and so on. But usually the motivational analyst does not offer to do this; he does sometimes apologize for his small number of depth interviews and suggests that perhaps 300 or 400 more would permit him to draw more meaningful conclusions. If this is done, it is usually only to compound the confusion because the results of 300 depth interviews

are simply a collection of that many atypical individual attitudes. Statistical quantification and standardized questioning offer the only means of evaluation.

QUANTITATIVE MOTIVATION ANALYSIS

We have found that motivation analysis does serve to uncover attitudes and motives which can then be incorporated in our standard research procedure; for example, we found at least fifteen reasons why certain individuals bought or didn't buy a given brand of beer. All of these reasons were then included as part of a standardized questionnaire. Interviews were conducted among a scientifically selected sample and statistical analysis of the interviews determined the relative importance of each of these reasons in a very succinct manner. Thus we found that a majority of regular drinkers didn't like a certain beer because it was too "sweet"; an almost equal proportion considered it too "hoppy." In the meantime, a competing brand was being bought by a proportion of regular drinkers because it was "dry," "not too sweet," "slightly bitter." This is the sort of information that the brewer can do something about both in modification of his product and also in his advertising.

In another recent case, as a result of unstructured interviews, we found a hint that the husband's attitude toward instant coffee might have an important influence on whether his wife bought it or not. When we came to quantify this, along with a number of other hints that were developed in the informal investigation, we found that it was indeed an important factor in the wife's purchase. We also found that in a majority of cases it was true that the less successful a housewife rated herself as a coffee maker, the more likely she was to use instant coffee. Conversely, we found that ease of preparation was of very little importance as an influencing motive in the purchase of instant coffee today. Everybody who buys instant coffee and everybody who does not buy it are equally aware of its convenience.

In other studies, motivation analysis has served to determine the effectiveness of advertising in creating new users for a product. In such a case, we first isolate all of the new users among the respondents, by statistical methods. Then we divide these into two groups, those who have been exposed to a given ad, and have an impression from it, and those who have not been exposed. The difference between these two groups gives us a key to the potency of the ads in creating new users for the product. This is similar to the procedure used in tests to

determine the value of the Salk polio vaccine and similar to methods used commonly in the ethical drug field to check the curative effects of a new drug.

No matter what may be the claims of the clinical psychologist attempting to apply his methods in marketing research, he might just as well recognize the impossibility of avoiding statistics. The clinical psychologist is very likely to attempt to draw conclusions before he draws graphs, but as has been stated by Professor Paul Meehl of the University of Minnesota in his book, *Clinical versus Statistical Prediction*, "Always, we might as well face it, the shadow of the statistician hovers in the background; always the actuary will have the final word."

PART II

Projective and Related Psychological Techniques

INTRODUCTION

As has been noted, the principal manifestation of psychological thinking in motivation research is through projective techniques. In fact, in many ways the manner and validity of the application of these projective techniques constitute the hub of the controversy, for in particular cases the argument revolves not around the use of psychological principles as such but rather around the validity of a specific technique in that particular application. Attention to the merits and limitations of projective techniques as applied to commercial research is therefore an important part of this volume.

It would seem only fitting to begin this section with a simple and concise description of the principal projective techniques used in commercial work with the advantages and disadvantages of each. Such an article is the one by L. O. Brown, which the editors feel fortunate to be able to reproduce in this volume. The piece by Steuart Britt neatly supplements this article by stressing four pitfalls frequently encountered in the use of psychological techniques for motivation research. The following two selections present specific examples of the use of projective techniques, the study by Mason Haire being perhaps the classic example of successful application and hence all the more useful because of the detail with which it is presented. Equally useful is the survey of applications of such techniques in the public opinion field by Weschler and Bernberg, which contains many implications to commercial research.

The articles so far having been largely favorable to projective techniques, it would seem useful to include two critical views at this point to provide marketing researchers with a better view of the other side of the picture. The articles by Lucas and Ferber are both of this type, concentrating on the difficulties of application of projective techniques to marketing problems.

The remainder of this part is devoted to technical discussions of the application of particular projective techniques. Written mainly by psychologists, the first three of these articles point up clearly the main problems in applying certain of these techniques, particularly in commercial research. Many researchers may be especially interested in the error-choice technique proposed by Hammond, both for its novelty and its provocative nature. The remaining four articles in this part

provide useful examples of the contention that, despite difficulties of application, when used properly projective techniques can contribute to solutions of marketing problems.

8. PRINCIPAL PSYCHOLOGICAL TECHNIQUES IN MOTIVATION RESEARCH*

Here is a more or less impartial treatment of the main psychological techniques, with emphasis on the advantages and disadvantages of each. As such, it is an apt introduction to the subject: the techniques are described simply and clearly, and the good and bad points of each are noted systematically.

Every individual has his own complex personality controlling his motivations. He interprets situations and adjusts his behavior to them in terms of his total past experience and temperament.

Personality, however, operates on at least three levels. First, there is the outer level which one shows to the world. When we speak of a person as being blustery or calm, we speak of this level. Second, there is the conscious inner level of personality. This level is the private world of wishful thinking, secret plans, and daydreaming. Third, there is the subconscious level. Here there are strong motivating forces which the individual himself does not recognize. All of us have had the experience of being surprised, or possibly shocked, by our own unexpected reactions to many situations. The basic problem of motivational research is to penetrate into the conscious and subconscious inner levels of personality.

A large variety of methods is employed in conducting motivation research. The procedures used in any specific project depend upon the training and background of the individuals conducting the research, as well as on the nature of the problem. A great deal of experimental work is also being conducted to develop more effective techniques. In the discussion which follows, the principal methods found most useful in motivation research are explained.

* Adapted from an article by Lyndon O. Brown, vice-president, Dancer-Fitzgerald-Sample, "What Motivational Research Is and How It Works; Its Advantages and Shortcomings," *Advertising Age*, Vol. 26, July 25, 1955, pp. 63 ff. This article is essentially Chapter 18 of the author's book, *Marketing and Distribution Research* (New York: Ronald Press, 1955).

DEPTH INTERVIEWING

The depth interview is an adaptation of the basic practice of clinical psychiatry. The psychiatrist places the respondent in the most relaxed and comfortable position possible, asks many questions, and makes statements to stimulate reactions. The psychiatrist works over a long period of time, usually from six to eight months in weekly sessions. Voluminous notes are made of the patient's answers to questions and free associations. The psychiatrist then spends a great deal of time studying these notes and other data obtained in the case history. The analysis is often supplemented by study of the results of various psychiatric tests that have been administered. Out of all this, he draws conclusions which provide a foundation for the therapy he will recommend.

In depth interviewing, a similar procedure is followed. It is vastly compressed in time, however, the typical depth interview with a consumer lasting from one to two hours. The common elements are the following: the work is done by a person trained in the techniques of probing, the interviewer asks questions he considers appropriate as the interview progresses, the questions are asked in an order developed during the interview, and the interviewer is alert to make observations of subtle reactions of the respondent. The questions are largely indirect in nature, but are centered around the product or problem involved.

A good example of the contrast between the depth interview and the ordinary questionnaire process is the following: The interviewer is attempting to determine the respondent's true feelings with respect to his position in a corporation. In a depth interview he casually mentions the name of one of the respondent's associates who has recently been promoted. He then observes reactions. After the respondent has indicated for the third time that he is not jealous of his associate's progress, the depth interviewer concludes that the truth is precisely the opposite, namely, that the respondent is motivated by deep feelings of jealousy. In the ordinary questionnaire process, a direct question would be asked, such as "How do you feel about X?"

Because all the principal basic motivations are rooted in fundamental and common psychological characteristics of people, a researcher employing the depth interviewing technique relies on a small number of interviews, generally from 50 to 200. After the interviews are completed, a psychologist with psychiatric training or a team of analysts studies the reports to determine the threads of common responses and reactions running through them. Applying their specialized knowl-

edge, a team of psychologists prepares an analytical report itemizing and evaluating the motivations on the basis of their psychological training, clinical experience, insight, and experience gained from previous work in marketing research.

The primary advantage of the depth interview technique is its ability to discover unsuspected new motivations. In the marketing operation, traditional patterns of selling and advertising appeals evolve and are generally followed. These appeals frequently depart from the motivations which really determine consumer behavior. Another consideration is that the relative importance of various motivating drives changes over a period of time. Through depth interviewing, one frequently discovers the strength of a new appeal so that an entirely different approach to the market can be made.

A second advantage of the depth interview procedure is that it lends itself effectively to the development of a motivational pattern with respect to the selection of a brand or other action under study. This pattern is particularly well structured in case the work has been done by people who have been trained in the comprehensive and systematic Freudian approach. Even though many of Freud's theories have no bearing on a marketing situation and some of his ideas have been greatly modified, the Freudian school has a relatively complete theory of human motivation, which relies basically on the depth interview, and people trained in this school are particularly skilled in developing a pattern of motivation. Once developed, this pattern is generally supported by logic and common sense.

A third advantage to the depth interview is that it provides a tremendous stimulus to the insight of the interviewer. In most marketing research based on interviewing, the knowledge gained is limited to that which comes from the respondent. In a depth interviewing study, so much of the final conclusion depends upon the beliefs and judgment of the people who do the interviewing that we have in effect a two-way street where both interviewer and interviewee are contributing to the knowledge gained. Experience with depth interviewing in marketing research has demonstrated time and again that the major contribution to the solution of the problem has often sprung from the clinical insight of the people conducting the research, rather than from the answers and observations recorded in the interviewer's notebooks.

There are a number of important weaknesses in the depth interviewing approach. The first is the tremendous compression of time the depth interview receives in contrast to the psychiatric interviewing

process on which it is based. While there has been repeated demonstration of the validity of psychiatric therapy, there is little or no proof that the depth interview as practiced in marketing research does actually define basic motivations with similar accuracy.

A second weakness of the depth interview in marketing research is the lack of a systematic structure for interpretation of the information obtained. Persons trained in the psychoanalytic school do have a systematic approach to the study of motivation. This approach, however, is with respect to the problems met in clinical psychiatry. No systematic approach has been developed which is specifically adapted to marketing research. Furthermore, the bulk of people who practice depth interviewing do not have the advantage of psychoanalytic training, hence have no systematic structure for interpretation.

Another difficulty is that no quantifiable data are obtained in the depth interviewing process. This means that human judgment is involved in summarizing the findings. Different results will often be obtained by different people in the same situation. As a result, there is little or no opportunity for verification.

A final weakness of depth interviewing is the lack of trained persons who are competent to do the interviewing work. The importance of this deficiency is exemplified in the case of one very successful organization practicing in this field. There is no doubt that the head of this organization can conduct extremely productive depth interviews himself. It is found, however, that other staff interviewers are unable to develop the same quality of penetrating information, in spite of the fact that they have studied extensively in the field of psychology. Depth interviewing is often a self-reflection of the one or two key individuals who have the primary responsibility for the project. It is a highly personal matter.

Furthermore, there are a great many people who purport to practice depth interviewing but who do not actually employ the technique as it is understood in the fields of psychiatry and psychology. It is fashionable to employ the term "depth interview." People without training conduct unstructured and fairly lengthy interviews. These are no more than an informal type of interviewing. Yet they are frequently advanced as true depth interviewing.

FOCUSED GROUP INTERVIEWING

In the focussed group interviewing method, a small number of representative consumers, usually from 6 to 12, is brought together for an informal group discussion of the matter under consideration.

The discussion leader is a trained psychologist experienced in group interviewing techniques. He has an outline of specific topics to provide the focus of the research. He allows the session to follow its natural course, however, bringing the discussion back to his topics as he feels the need to do so.

In addition to subjects for discussion, the group leader may also introduce various stimuli, such as products, packages, pictures, and advertisements. The entire discussion is recorded on a wire or tape recorder. Comparisons of recordings of one session with those of another often show great contrast. After sessions with several groups have been conducted, the discussion leader listens repeatedly to the recordings. He then analyzes their content against the background of his psychological training and draws his conclusions with respect to the motivations uncovered.

Some similarities may be seen between the focussed group interviewing procedure and the depth interviewing technique. The order of the subjects discussed in the group interview and the form of questions are completely unstructured. A large amount of time is given to probing and stimulating questions. A typical group session, for example, lasts about three hours. Reactions to various stimuli introduced are observed carefully.

There are fundamental contrasts, however, between focussed group interviewing and depth interviewing. The principal contrast is that group interviewing depends primarily on the interaction of ideas, attitudes, emotions, and beliefs among the various members of the group. The theory is that motivations rise to the surface most effectively in the process of action and counteraction produced by various statements made by group members. For example, when one woman says, "I find soap is terribly hard on my hands," other members of the group are bound to respond strongly, to expand this thought considerably, and to disclose their own reactions. In the depth interview, we rely upon one individual for this response, expansion, and reaction. In the group interview, we rely on discussion and often obtain heated debate. This process is called social facilitation.

The advantages generally claimed for the focussed group method are:

1. Interstimulation broadens the base of communication. People in groups share ideas—more points emerge to think about, discuss, and evaluate than is true of a two-person conversation.

2. The threshold for personal revelations is lowered. One person makes a "daring" or intolerant statement which is accepted by the others, someone else

feels encouraged to speak more personally, and so on, until the participants as a group move toward a mutually sanctioned standard of frankness which one respondent talking privately to an interviewer cannot achieve so readily.

3. Social inhibitions cause individuals to respond more nearly in line with their behavior. The exaggerator, the fabricator, the irresponsible respondent is considerably "toned down" when talking for group consumption.

4. Respondents in groups try hard to contribute. *Esprit de corps* usually develops, which means that a high percentage of the respondents are motivated to give as much information, to express as many views, as possible.

5. Note-taking is facilitated. The tape recordings provide a reservoir of realistic communication which may be studied from all points of view. Recordings catch the "flavor" of words used by participants—the emphasis and nuances of oral expressions.

PROJECTIVE TECHNIQUES

One of the most interesting developments in motivation research is the use of projective techniques for penetrating the conscious and subconscious motivations of human behavior. Projective devices operate on the principle of confronting an individual with a purposely ambiguous situation which he must interpret and structure. The ambiguous situation may just be a word, such as *candy*, an incomplete sentence, or a picture. The subject is simply required to respond to the materials in some manner which seems appropriate to him. For example, he may be asked to tell a story about a picture. Within the framework of such flexible directions, complete freedom is allowed. In his response to such stimuli, a person discloses his private world of attitudes, feelings, and values. The projective devices tend to remove the inhibitions of the individual because he is responding in terms of other people rather than thinking of himself.

An illustration of the basic principle of the projective technique in very simple form is found in a slight change in wording of a preference question. Housewives were shown two statements, each of which described a laundry product in a different way and asked, "Which of these two products would you prefer to use?" Their responses to this direct, subjective questioning were 20 per cent in favor of statement A and 80 per cent in favor of statement B. A comparable group of housewives was shown the identical statements but with only a very slight change in the wording of the question to, "Which of these two products *do you believe most women* would prefer to use?" Their responses to this indirect, projective questioning were 53 per cent in favor of statement A and 47 per cent in favor of statement B.

It is seen that this slight change in the wording of a question actually reversed the preference vote. The nature of the change was to

shift the frame of reference from the respondent herself to an outside situation (other women); hence, in principle the second question form was an application of the projective technique. When confronted with the direct question, housewives would not admit a preference for statement A; by the projective device a large proportion displayed their actual preference for that statement.

Projective techniques are designed to tap the deeper strata of personality in terms of unconscious motivation. In any direct question situation, even the most co-operative subject who has no intent to falsify can tell only what he knows about himself. It is generally accepted that we are inclined to deny many of our true motives and to rationalize impulses in order to make our behavior more socially and personally acceptable. The material brought forth by means of projective methods goes beyond these superficial defenses. For example, it is often very difficult to get a true evaluation of housewives' attitudes toward various home service personalities because of women's apathy or reluctance to say unfavorable things about them.

Projective tests make use of what a person selects for response to a planned stimulus and of his characteristic manner in organizing and deriving meaning from the ambiguous field confronting him. How a person responds to an idea or condition depends upon his own needs, wishes, and preoccupations.

There are several dangers in the use of projective techniques. They must be administered and interpreted by trained, experienced practitioners. Since the tests are largely unstructured, their administration by unskilled persons can cause serious biases. The interpretation of the results can be misleading or meaningless unless undertaken by those who know what to look for and how to do so.

The key to the successful use of projective devices lies in three areas. The first requirement is the development and selection of materials (stimuli) which are most successful in obtaining responses that project beyond the conscious surface level. The second requirement is that the stimuli evoke responses which are related to the motivational problem being researched. The third requirement is that the responses can be interpreted into a logical motivational pattern.

Free Word Association. Word association is one of the oldest of projective techniques. The respondent is read a list of words, one at a time, and asked to respond with the first word that comes to mind. In this way, by careful planning of the order in which the test words appear, the replies become almost automatic.

The data-gathering procedure for word association is standardized.

The investigator makes a preliminary statement such as the following to the respondent: "I am making a study of the way different words are put together in people's minds. I will read you a list of words one at a time, and I want you to tell me the very next word you think of. Any word is all right—the main thing is speed. For example, if I said *paper*, you might say *book*; do you understand?"

The interviewer then reads a list of words carefully selected to reveal motivations with respect to the subject of the research and records word responses.

Word association data are usually judged in three ways: by frequency with which any word is given as a response, by analysis of the amount of time that elapses before a response word is given to a test word (hesitation), and by the number of respondents who cannot give any response at all to the test word after a reasonable period of time (blocking).

Common Responses. An analysis of the frequency with which a particular word is given as a response is an analysis of the common responses. Although it might seem that the unconscious attitudes of an individual would be so different from those of any other individual that no two could be grouped for analysis, the opposite is usually the case. When unconscious attitudes are tapped, they are strikingly similar for most people. Thus, the number of times that a word is given as a response to a test word shows a basic attitude toward the test word.

Hesitation. The respondent is timed from the moment the test word is spoken until he gives a response. In laboratory work, a calibrated stop watch is used. In home interviews, investigators count seconds for recording hesitations, after practicing their estimates with a watch. When the respondent takes more than three seconds to respond, it is called a "hesitation." A hesitation indicates that the respondent is involved in some way with the test word or what it symbolizes; his response is not immediate and automatic. He may, for example, be substituting a second response for his immediate reaction, which he feels is not acceptable. He is able to respond to the word in some fashion, however, even though he hesitates. Hesitation shows the comparative emotional involvement for each word.

Nonresponse. In some cases the test word or what it symbolizes is so charged with emotion for the respondent that he "blocks" so that he can give no response at all. It has been demonstrated that, for such words, little if any consistent communication is possible. These words simply do not convey a conscious message and are charged with strong emotional inhibition.

These methods must be used in combination for anything more than the most simple word association analysis. Among the most useful combinations in indicating emotional involvement with a word is the combination of the "hesitation" and "no-response" rates. For example, if a word shows a high hesitation rate and a low no-response rate, we can assume that, although there is concern about the area that the word symbolizes (as shown by the high hesitation rate), the low no-response rate shows that the word does convey its message. This type of response often indicates the attention-attracting word for conveying a message. When a word has both a high hesitation rate and a high no-response rate, however, the word is usually a bad word for conveying a message. When a word or group of words with high hesitation and no-response rates are presented in a message, the meaning of the message is usually highly distorted in its transmission or simply not transmitted at all.

Successive Word Association. This technique is similar to single-word association except that the respondent is asked to continue to give single-word responses to the same test word as long as he can. As single-word association gives the most common response, successive word association continues through several levels of attitudes. This method shows which words and attitudes are in the process of becoming common and indicates the growth of favorable or unfavorable attitudes. Successive word associations are judged in ways similar to single-word associations with the aid of common responses and no-response rates.

Sentence Completion. Closely akin to word association is the sentence completion technique. In this method the respondent is presented with a series of statements which express incomplete ideas but direct his attention along a predetermined path. The respondent is asked to complete each of the sentences with whatever thoughts first come to mind. Thus, he finishes each sentence with his own interpretation of what the beginning means to him.

The sentence completion method operates on the principle that the associations revealed by responses to the incomplete stimuli reveal hidden motivations. In the construction of the incomplete sentences the researcher deliberately varies the nature of the ideas presented so that the respondent is not aware of the specific sentences which are inserted to obtain data directly related to the problem at hand. The sentences are kept as simple as possible to encourage rapid completion so that the the respondent will shift his thoughts back and forth between different subjects as he completes the series of sentences.

Sentence completion is used primarily in conjunction with other projective methods to amplify the data obtained by other means. The

results produced by this technique are also valuable as a basis for verification of interpretations and conclusions.

Picture Responses (Frustration). In the picture response technique, the respondent is shown an illustration which may be interpreted in many different ways and is asked to interpret or tell its story. The pictures are purposely made quite vague in order to provide the greatest possible opportunity for a variety of responses. The Rorschach ink-blot test is used in clinical psychology. These "pictures" are literally made by ink blots so that only an apparently meaningless form is shown. It has been found, however, that when people are asked to describe what such "pictures" convey, they do project their own personality traits into the pictures. The Thematic Apperception Test presents a subject with a series of social situations which he tries to explain, tells how the situation came about, and describes the outcome of it.

An example of the picture response technique, based on an adaptation of the Thematic Apperception Test, is afforded in a study of motivations with respect to automobiles made by Social Research Incorporated. They used a picture of the road seen from behind the steering wheel position. Only the hands grasping the wheel were shown, the respondent was told to put himself in that position, and asked to tell what kind of things came to his mind. The speedometer was shown in the picture—set between 60 and 70 miles per hour.

Although the instructions told the person to put himself in the driver's situation, most people did not do this because they did not want to think of themselves as going so fast. The picture clearly elicited contradictory attitudes about speed. Respondents told stories about how much fun it would be going this fast somewhere on the open road, but it was always someone else doing it; he should not be doing it. Social Research translated this to mean that a car must be capable of high speed, but high speed is wrong. Speed involves both pleasure and fear. The manufacturer must therefore convince prospects that his automobile is both fast and safe.

In marketing research, the picture response technique usually employs illustrations of situations involving the use of the product or service under study. One of the most effective forms is the picture cartoon. In this device the picture usually shows two persons. One of them is making a statement in the speech balloon. An example would be a picture of two housewives talking over the back-yard fence. The first housewife is saying, "I just served my family Swift's new precooked sausages." The balloon for the second housewife's response is left blank and the respondent is asked to tell what she said. Psychologists

interpret these responses in terms of the attitudes, reactions, doubts, fears, and other motivations revealed with respect to this new product.

Experience in marketing research with picture response devices clearly indicates that they succeed in getting people to respond impersonally to the situations confronting them. There is no question that these responses are quite different from those which are obtained by direct questioning. Also, there is little doubt that this device gets much closer to inner motivations.

One of the advantages of the picture response method is the control of the researcher over the situations presented to the respondents. It is possible to explore many different aspects of motivation and to relate them rather directly to the specific problem.

Another advantage of this method is that it is particularly applicable to situations in which the respondent is most likely to hide his real feelings because of social pressures. An example is found in a study of motivations for enlisting in the army. It is not socially acceptable for one to give avoidance of the draft as a reason impelling enlistment. In response to a picture cartoon, however, which confronts the respondent with the statement, "I hear Bob has enlisted. Why do you suppose he did that?" it is perfectly proper for a young man to ascribe the draft as a motivating force for an anonymous Bob.

Another important advantage of the picture response technique is that a relatively large number of respondents can be covered economically in the research, and responses to a variety of situations can be obtained from each one. The responses also tend to pattern themselves. As a result, analysis into basic motivation themes running through responses from different subjects is facilitated.

LIMITATIONS OF MOTIVATION RESEARCH

As more and more motivation work is being done, leaders in the field are raising questions as to the precise scope which should be assigned to it. Because of the intense desire to obtain sound answers as to basic motivations, there is a natural tendency to assume the existence of some single dominating motivation or drive behind any given form of market behavior. There is a growing feeling, however, that there is no such thing as one most important motivation in the market place, but rather that there is a variety of motivations toward which marketing efforts may be directed. Furthermore, researchers are coming to the point of view that the differences in the power of the drives represented by a number of motivations are not as great as might be assumed. A number of leading practitioners now believe that moti-

vation research is primarily a verbalizing device which helps make respondents more articulate in expressing their reasons for buying a product or selecting one brand over another, rather than a mystical procedure for discovering some one all-embracing hidden motivation.

There is also a growing caution against automatic acceptance of motivation devices developed in other fields. These devices have been individual-centered, dealing with a single person, rather than mass-centered. The problem of the psychiatrist is to develop a tremendous mass of data covering the entire life of an individual. Out of this mass of data he arrives at certain conclusions regarding factors motivating the patient's current behavior. In applying motivation techniques to the masses of individuals who compose a market, we must work with a relatively small amount of data regarding each individual who is studied. These relatively thin data on many individuals must then be combined into a general rationalization regarding large social groups. Whether this is a legitimate transfer is a moot question.

Related to the preceding difficulty is the question as to whether motivation data on individuals are even additive. In a market survey, we can learn facts about the behavior of individuals with respect to phenomena, such as brand preferences, and add them up to arrive at conclusions regarding the total market and its various subdivisions. It is doubtful whether the motivations compelling one individual toward a certain type of behavior are sufficiently homogeneous so that we can add them together.

A fourth concern regarding motivation research is that the data are not usually quantified. Since quantification is generally regarded as an essential element in scientific method, its absence in motivation research removes a key process creating confidence in a result. There is always room for doubt as to the true significance of any given finding. This lack of quantification also leads to questions regarding sampling structure. While it is generally accepted that motivation research can work effectively with small samples, just how one is to determine how many and which individuals are to be included in a motivation study is largely a matter of opinion.

A fifth problem is that of interpretation. Since the results are not expressed in quantitative form, interpretation becomes a personal matter. This means that the results of a motivation study largely reflect the frame of reference of the individual who is responsible for its interpretation. Obviously, his own motivations, training, attitudes, beliefs, and desires constantly tend to influence the results. In order to minimize this difficulty it is a standard practice to have the field find-

ings analyzed and interpreted independently by two or more individuals. If their interpretations coincide, it is assumed that the conclusions are correct. Analyses and interpretations, however, should always be in the hands of trained social psychologists with a great deal of experience in this area. Unfortunately, much of this work is done by amateurs and laymen who do not have the advantage of training, experience, and the resulting objective and cautious point of view which is essential to sound interpretation of motivation data.

Motivation research is also plagued by major problems of semantics. In developing its own vocabulary, terminologies have arisen which tend to be complicated. Worst of all, there are many different schools of thought in the various disciplines represented in this field; many of them have their own unique terminologies and, at the same time, use the same words with different meanings. A common difficulty in all the social sciences is a lack of precise symbols such as those employed in chemistry, physics, and mathematics. Furthermore, instead of using completely coined words or a dead language such as the Latin used in medicine and law, many of their words are similar to those in common English usage. This leads to vagueness and considerable difficulty in communication.

A seventh major problem is the lack of positive validation of conclusions. The closest approach to validation has been verification by obtaining similar interpretations of parallel situations by different analysts. Motivation research therefore suffers from the lack of any outside criteria which can be employed to validate its conclusions. As a result, the question is frequently raised, "Does this test measure what it sets out to measure?" The fact that many of the methods produce clear-cut behavior responses does not necessarily mean that these responses are to be interpreted as a direct reflection of the motivations and drives they purport to measure. All motivation research makes the assumption that there is a direct and definable relationship between what is observed in the research and the motivation the researcher is seeking.

In spite of the difficulties discussed above, there is no doubt that the current techniques of motivation research have taken us a long way down the tortuous road toward understanding human behavior. The actual experience in its applications to marketing problems is still quite limited. As more experience is obtained, many of these questions will be answered and the failures overcome by more effective and dependable methodology.

9. SOME HAZARDS OF MOTIVATION TECHNIQUES*

The four hazards discussed in this brief article are applicable to almost any set of research techniques. That they are particularly applicable to psychological techniques as applied to motivation research is perhaps the central point of this article, and there is no doubt that this point is well taken.

If you are considering the use of motivation research: (1) be sure you really understand the term; (2) don't expect it to solve all problems; (3) ask for evidence on the methods proposed; and (4) make certain you use valid techniques.

Actually, motivation research is not new. It has been part of the general body of scientific knowledge for many years. Agencies have been doing this sort of work for some time. Psychological techniques have been applied to the marketing problems of many companies. And the techniques that have been employed—which now might have a "motivation research" label put on them—have involved depth interviewing, group interviewing, narrative probes, picture probes, sentence-completion methods, word-association techniques, and so on.

We have found that these techniques used with other marketing research techniques can be very helpful. The danger is in thinking that these psychological techniques will always produce conclusive answers to any and all problems.

There are four hazards in motivation research: (1) the term is misleading, (2) it is no panacea, (3) not all of its technics are reliable, and (4) not all of its technics are valid.

Term "Motivation" Misleading. Those who carry the torch for motivation research to the exclusion of all other kinds tend to give the impression that all other kinds of research do not involve a study of motivation. Yet marketing and advertising personnel and salespeople have been concerned with man's motives for a long time, and they will continue to be interested in them. The man in the street is also concerned with the reasons why he does what he does, but he probably has not learned the lingo of motivation research.

Certain psychological or clinical techniques should not be given the

* Adapted from an article by Steuart Henderson Britt, Northwestern University, and Managing Editor, *Journal of Marketing*, "Four Hazards of Motivation Research: How to Avoid Them," *Printers' Ink*, Vol. 251 (June 17, 1955), pp. 40 ff.

fancy name of motivation research. Why should the name of a technique be given the name of the objective it is supposed to accomplish? And certainly it is a mistake to think that the psychological techniques are the only tools for learning about people's motives.

Motivation Research No Panacea. The techniques are not new. They have been known for many years to psychologists and various social scientists. The point is that these techniques, invented for the most part by academic psychologists, have now been discovered by men in the market place and are being widely exploited. Downright suspicion by practical businessmen of a few years ago has now given way to an attitude of joining up as soon as possible. It seems that almost everyone is applying for membership in the M. R. Society, with all of its mysteries.

We find that straight consumer surveys will often yield answers that cannot be obtained from psychological techniques. Vice versa, psychological methods can pick up information that is not always obtained through the usual consumer surveys. Or, to put it another way, psychological techniques often will provide clues that can be followed up with detailed questionnaire surveys to find out the extent to which the clues are correct.

This can also work the other way around. We may find through the more formalized questionnaires that people have certain attitudes. Then, through projective techniques and other psychological methods, these attitudes can be probed more deeply.

There is room for both kinds of research. The kind of problem determines the extent to which each is used—either individually or in connection with other techniques.

Not All Techniques Are Reliable. In explaining why all techniques are not reliable, we shall have to refer to the statistical concept of reliability. In elementary statistics, the word "reliability" refers to the precision with which a given test or technique measures a function. For example, if you take your temperature each day, you want a thermometer that is reliable: one that will consistently give the same measurement for the same set of conditions. The problems encountered in the use of projective techniques in marketing are also ones of reliability. What evidence is there that projective techniques used by some organizations are as reliable as they should be?

Another way to put this is, how do we know that, if the same research techniques were used on the same people or a comparable sample, we would get essentially the same results? This kind of reliability has been demonstrated for a good deal of quantitative consumer re-

search. It needs to be demonstrated by certain organizations with their projective techniques.

Related to this is the question: What about the reliability of the judgments of those who interpret the materials? It should be emphasized that this is quite a different thing from the honesty of the judgments. This is not a question of sincerity of purpose or honesty of intent. It is a question of consistency of interpretation from one person to another. Would the same interpreter of the materials make the same interpretation if given the same set of data two weeks later? Or would a second interpreter of the materials give the same interpretation as the first one?

Many years ago a well-known social scientist, Stuart Rice, reported information concerning interview records of social workers who had been engaged in interviewing 2,000 homeless men who were applying for free lodging. Since the average interviewing form used by social workers is well standardized, you might assume that the answers filled out would be completely objective. Yet Rice showed how the already formed attitudes of the social workers might affect their behavior in filling out these forms.

One of the interviewers was a strong prohibitionist, another was an ardent socialist. In analyzing the interview blanks turned in by the prohibitionist and socialist, Rice found that the prohibitionist attributed 62 per cent of the cases to liquor and only 7 per cent to industrial conditions. On the other hand, the socialist attributed 39 per cent to industrial conditions and only 22 per cent to liquor. According to the information turned in by the prohibitionist, 34 per cent of the applicants had mentioned liquor as a reason for their downfall, whereas the socialist reported that 60 per cent had mentioned industrial conditions as the principal reason!

This study is an excellent illustration of the way that already acquired attitudes may affect what is done in a supposedly objective manner. Applying this analogy to the area of projective techniques, the question is: To what extent will two different interpreters of projective materials give the same interpretation of the same materials? Or, better still, to what extent are two different practitioners of motivation research likely to give the same interpretation to projective data?

Validity of Techniques. In statistics, "validity" is not the same thing as "reliability." Validity means that something measures what it is supposed to measure. Again—using the example of the clinical thermometer—the thermometer should be one that measures temperature, not blood pressure or pulse rate. How do we know that the pro-

jective techniques have validity? How do we know that the techniques measure what they are supposed to measure?

In psychology and psychometrics we have a phrase, "face validity." The term "face validity" means that something on its face appears or seems to have validity. It seems to measure what it is supposed to measure—although it may turn out that it really does not have very much validity.

There are numerous examples of this from the field of automobile-driver training. Dozens of driver tests have been invented where people sit in make-believe automobiles, push pedals, react to flashing lights, turn the steering wheel and so on—and measures are obtained of their motor proficiency. Some of these driver tests eventually have been demonstrated to have some validity (to help sort out those people who will make good drivers under actual road conditions as compared with those who will not). But many of these gadget-like tests turn out to have just face validity. They look good, but they just don't test out.

Applying this idea to the projective techniques, we should take a good look to see if the techniques have only face validity, or if they turn out to be accurate measures of what they are supposed to measure. There is a lot of evidence that some of the techniques under certain conditions do an extremely good job. But more evidence is needed. The only way to find out is to set up a number of actual experiments in this highly complex field.

SUMMARY

There is a lot to be learned about the "whys" of consumer actions and habits from the use of the psychological techniques that have been dubbed "motivation research." They have to be used in the right places, however, and at the right times. Certainly the more conventional approach of the structured, precoded questionnaire is still one of our most useful research tools.

Here is how the four difficulties of motivation research may be avoided:

1. The next time you hear motivation research discussed, just remember that this phrase really refers to certain psychological measuring techniques. Obviously, there is no way to stem the tide of using the popular phrase, "motivation research." But please remember what the phrase really means. And remember that lots of other research work also involves a study of motivation.

2. Keep in mind that motivation research is not a cure-all. The techniques are useful under certain circumstances. But they should not

be seized upon as the solution to *all* marketing problems just because they seem to be the latest thing.

3. When someone talks to you about his special motivation research techniques, ask him to demonstrate the reliability of his techniques. Ask for experimental evidence on this, and also on the reliability of the *interpretation* of the materials.

One solution to this is to use only techniques whose reliability has been established, and then to check on the consistency of interpretation by having two or more different individuals do the interpreting of the results. These analysts of the projective results should work separately and independently, and their reports can then be gone over and discussed by a third analyst.

4. Ask about the validity of the projective methods proposed. Better still, try to set up actual tests in your own organization that will show whether the projective results have truly demonstrated what they are supposed to demonstrate. The projective methods can well be employed, but careful testing is needed for each marketing situation. After all, no two marketing or sales situations are ever the same.

10. AN APPLICATION OF PROJECTIVE TECHNIQUES*

This example of the successful application of projective techniques to marketing research is well on the way to becoming a classic in the field. It is a well-deserved accolade, and contrary to other "classics" is very clearly written, with the research techniques spelled out in considerable detail. Of particular interest in the latter respect is the manner in which statistical sampling analysis is used to supplement the projective findings and how even the results of the projective techniques are placed on a quantitative basis.

It is a well-accepted maxim in merchandising that, in many areas, we are selling the sizzle rather than the steak. Our market research techniques, however, in many of these same areas, are directed toward the steak. The sizzle is the subjective reaction of the consumer; the steak the objective characteristics of the product. The consumer's behavior will be based on the former rather than the latter set of characteristics. How can we come to know them better?

* Adapted from an article by Mason Haire, University of California, "Projective Techniques in Marketing Research," *Journal of Marketing*, Vol. 14 (April, 1950), pp. 649-56.

When we approach a consumer directly with questions about his reaction to a product we often get false and misleading answers. Very often this is because the question which we heard ourselves ask was not the one (or not the only one) that the respondent heard. For example: A brewery made two kinds of beer. To guide their merchandising techniques they wanted to know what kind of people drank each kind, and particularly, what differences there were between the two groups of consumers. A survey was conducted which led up to the questions, "Do you drink beer?" (If *yes*) "Do you drink the Light or Regular?" (These were the two trade names under which the company marketed.) After identifying the consumers of each product, it was possible to find out about the characteristics of each group so that appropriate appeals could be used, media chosen, and so forth.

An interesting anomaly appeared in the survey data, however. The interviewing showed (on a reliable sample) that consumers drank Light over Regular in the ratio of 3 to 1. The company had been producing and selling Regular over Light for some time in a ratio of 9 to 1. Clearly, the attempt to identify characteristics of the two kinds was a failure. What made it miss so far?

When we say "Do you drink Light or Regular?" we are at once asking which brand is used, but also, to some extent, saying "Do you drink the regular run-of-the-mill product or do you drink the one that is more refined and shows more discrimination and taste?" The preponderance of "Light" undoubtedly flows from this kind of distortion.

When we ask questions of this sort about the product, we are very often asking also about the respondent. Not only do we say, "What is product like?" but, indirectly, "What are you like?" Our responses are often made up of both elements inextricably interwoven. The answers to the second question will carry clichés and stereotypes, blocks, inhibitions, and distortions, whenever we approach an area that challenges the person's idea of himself.

There are many things that we need to know about a consumer's reaction to a product that he cannot tell us because they are to some extent socially unacceptable. For instance, the snob appeal of a product vitally influences its sale, but it is a thing that the consumer will not like to discuss explicitly. In other cases the consumer is influenced by motives of which he is, perhaps, vaguely aware, but which he finds difficult to put into words. The interviewer-respondent relationship puts a good deal of pressure on him to reply and to make sense in his reply. Consequently, he gives us stereotypical responses that use clichés which are commonly acceptable but do not necessarily represent the

true motives. Many of our motives do not, in fact, "make sense," and are not logical. The question-answer relation demands sense above all. If the response does not represent the true state of affairs, the interviewer will never know it. He will go away. If it does not make sense it may represent the truth, but the respondent will feel like a fool and the interviewer will not go away. Much better produce a cliché and be rid of him.

THE NATURE OF PROJECTIVE TESTS

Still other kinds of motives exist of which the respondent may not be explicitly conscious himself. The product may be seen by him as related to things or people or values in his life, or as having a certain role in the scheme of things, and yet he may be quite unable, in response to a direct question, to describe these aspects of the object. Nevertheless, these characteristics may be of great importance as motives. How can we get at them?

Clinical psychologists have long been faced with a parallel set of problems. It is quite usual for a patient to be unable or unwilling to tell the therapist directly what kinds of things are stirring in his motivational pattern. Information about these drives are of vital importance to the process of cure, so a good deal of research has been directed toward the development of techniques to identify and define them. The development of projective techniques as diagnostic tools has provided one of the most useful means to uncover such motivations, and the market-researcher can well afford to borrow their essentials from the therapist.

Basically, a projective test involves presenting the subject with an ambiguous stimulus—one that does not quite make sense in itself—and asking him to make sense of it. The theory is that in order for it to make sense he will have to add to it—to fill out the picture—and in so doing he projects part of himself into it. Since we know what was in the original stimulus, we can quite easily identify the parts that were added, and, in this way, painlessly obtain information about the person.

Examples of these tests come readily to hand. Nearly everyone is familiar with the Rorschach Test, in which a subject is shown a series of ink-blot and asked to tell what they look like. Here the stimulus is incomplete in itself, and the interpretation supplied by the patient provides useful information. This test yields fairly general answers about the personality, however, and often we would like to narrow down the area in which the patient is supplying information.

The Thematic Apperception Test offers a good example of this

function. Let us suppose that with a particular patient we have reason to suppose that his relation to figures of authority is crucial to his therapeutic problem. We can give him a series of pictures where people are shown, but where the relationship of authority or the characteristics of the authoritarian figure are not complete. He is asked to tell a story about each picture. If in each story the subordinate finally kills the figure of authority, we have certain kinds of knowledge; if, on the other hand, he always builds the story so that the subordinate figure achieves a secure and comfortable dependence, we have quite different information. It is often quite impossible to get the subject to tell us these things directly. Either he cannot or will not do so. Indirectly, however, he will tell us how he sees authority. Can we get him, similarly, to tell us how a product looks to him in his private view of the world?

A PROJECTIVE TEST IN MARKET RESEARCH

Let us look at such an example applied to market research. For the purposes of experiment a conventional survey was made of attitudes toward Nescafé, an instant coffee. The questionnaire included the questions, "Do you use instant coffee?" (If *No*) "What do you dislike about it?" The bulk of the unfavorable responses fell into the general area "I don't like the flavor." This is such an easy answer to a complex question that one may suspect it is a stereotype, which at once gives a sensible response to get rid of the interviewer and conceals other motives. How can we get behind this facade?

In this case an indirect approach was used. Two shopping lists were prepared. They were identical in all respects, except that one list specified Nescafé and one Maxwell House Coffee. They were administered to alternate subjects, with no subject knowing of the existence of the other list. The instructions were: "Read the shopping list below. Try to project yourself into the situation as far as possible until you can more or less characterize the woman who bought the groceries. Then write a brief description of her personality and character. Wherever possible indicate what factors influenced your judgment."

Shopping List I

Pound and a half of hamburger
2 loaves Wonder bread
bunch of carrots
1 can Rumford's Baking Powder
Nescafé instant coffee
2 cans Del Monte peaches
5 lbs. potatoes

Shopping List II

Pound and a half of hamburger
2 loaves Wonder bread
bunch of carrots
1 can Rumford's Baking Powder
1 lb. Maxwell House Coffee (Drip Grind)
2 cans Del Monte peaches
5 lbs. potatoes

Fifty people responded to each of the two shopping lists given above. The responses to these shopping lists provided some very interesting material. The following main characteristics of their descriptions can be given:

1. 48 per cent of the people described the woman who bought Nescafé as lazy; 4 per cent described the woman who bought Maxwell House as lazy.
2. 48 per cent of the people described the woman who bought Nescafé as failing to plan household purchases and schedules well; 12 per cent described the woman who bought Maxwell House this way.
3. 4 per cent described the Nescafé woman as thrifty; 16 per cent described the Maxwell House woman as thrifty. 12 per cent described the Nescafé woman as spend-thrift; 0 per cent described the Maxwell House woman this way.
4. 16 per cent described the Nescafé woman as not a good wife; 0 per cent described the Maxwell House woman this way. 4 per cent described the Nescafé woman as a good wife; 16 per cent described the Maxwell House woman as a good wife.

A clear picture begins to form here. Instant coffee represents a departure from "home-made" coffee, and the traditions with respect to caring for one's family. Coffee-making is taken seriously, with vigorous proponents for laborious drip and filter-paper methods, firm believers in coffee boiled in a battered sauce pan, and the like. Coffee drinking is a form of intimacy and relaxation that gives it a special character.

On the one hand, coffee making is an art. It is quite common to hear a woman say, "I can't seem to make good coffee," in the same way that one might say, "I can't learn to play the violin." It is acceptable to confess this inadequacy, for making coffee well is a mysterious touch that belongs, in a shadowy tradition, to the plump, aproned figure who is a little lost outside her kitchen but who has a sure sense in it and among its tools.

Coffee has a peculiar role in relation to the household and the home-and-family character. We may well have a picture, in the background, of a big black range that is always hot with baking and cooking, and has a big enamelled pot of coffee warming at the back. When a neigh-

bor drops in during the morning, a cup of coffee is a medium of hospitality that does somewhat the same thing as cocktails in the late afternoon, but does it in a broader sphere.

These are real and important aspects of coffee. They are not physical characteristics of the product, but they are real values in the consumer's life, and they influence his purchasing. We need to know and assess them. The "labor-saving" aspect of instant coffee, far from being an asset, may be a liability in that it violates these traditions. How often have we heard a wife respond to "This cake is delicious!" with a pretty blush and "Thank you—I made it with such-and-such a prepared cake mix." This response is so invariable as to seem almost compulsive. It is almost unthinkable to anticipate a reply, "Thank you, I made it with Pillsbury's flour and Borden's milk." Here the specifications are unnecessary. All that is relevant is the implied "I made it"—the art and the credit are carried directly by the verb that covers the process of mixing and processing the ingredients. In ready-mixed foods there seems to be a compulsive drive to refuse credit for the product, because the accomplishment is not the housewife's but the company's.

In this experiment, as a penalty for using "synthetics," the woman who buys Nescafé pays the price of being seen as lazy, spendthrift, a poor wife, and as failing to plan well for her family. The people who rejected instant coffee in the original direct question blamed its flavor. We may well wonder if their dislike of instant coffee was not to a large extent occasioned by a fear of being seen by one's self and others in the role they projected onto the Nescafé woman in the description. When asked directly, however, it is difficult to respond with this. One cannot say, "I don't use Nescafé because people will think I am lazy and not a good wife." Yet we know from these data that the feeling regarding laziness and shiftlessness was there. Later studies (reported below) showed that it determined buying habits, and that something could be done about it.

Analysis of Responses. Some examples of the type of response received will show the kind of material obtained and how it may be analyzed. Three examples of each group are given below.

*Descriptions of a woman who bought, among
other things, Maxwell House Coffee*

I'd say she was a practical, frugal woman. She bought too many potatoes. She must like to cook and bake as she included baking powder. She must not care much about her figure as she does not discriminate about the food she buys.

The woman is quite influenced by advertising as she signified by the specific

name brands on her shopping list. She probably is quite set in her ways and accepts no substitutes.

I have been able to observe several hundred women shoppers who have made very similar purchases to that listed above, and the only clue that I can detect that may have some bearing on her personality is the Del Monte peaches. This item when purchased singly indicates that she may be anxious to please either herself or members of her family with a "treat." She is probably a thrifty, sensible housewife.

*Descriptions of a woman who bought, among
other things, Nescafé Instant Coffee*

This woman appears to be either single or living alone. I would guess that she had an office job. Apparently, she likes to sleep late in the morning, basing my assumption on what she bought such as Instant Coffee which can be made in a hurry. She probably also has can [sic] peaches for breakfast, cans being easy to open. Assuming that she is just average, as opposed to those dazzling natural beauties who do not need much time to make up, she must appear rather sloppy, taking little time to make up in the morning. She is also used to eating supper out, too. Perhaps alone rather than with an escort. An old maid probably.

She seems to be lazy, because of her purchases of canned peaches and instant coffee. She doesn't seem to think, because she bought two loaves of bread, and then baking powder, unless she's thinking of making a cake. She probably just got married.

I think the woman is the type who never thinks ahead very far—the type who always sends Junior to the store to buy one item at a time. Also she is fundamentally lazy. All the items, with possible exception of the Rumford's, are easily prepared items. The girl may be an office girl who is just living from one day to the next in a sort of haphazard sort of life.

As we read these complete responses we begin to get a feeling for the picture that is created by Nescafé. It is particularly interesting to notice that the Nescafé woman is protected, to some extent, from the opprobrium of being lazy and haphazard by being seen as a single "office girl"—a role that relieves one from guilt for not being interested in the home and food preparation.

The references to peaches are significant. In one case (Maxwell House) they are singled out as a sign that the woman is thoughtfully preparing a "treat" for her family. On the other hand, when the Nescafé woman buys them it is evidence that she is lazy, since their "canned" character is seen as central.

In terms of the sort of results presented above, it may be useful to demonstrate the way these stories are coded. The following items are extracted from the six stories quoted:

<i>Maxwell House</i>	<i>Nescafé</i>
1. Practical Frugal Likes to cook	1. Single Office girl Sloppy Old maid
2. Influenced by advertising Set in her ways	2. Lazy Does not plan Newlywed
3. Interested in family Thrifty Sensible	3. Lazy Does not plan Office girl

Items such as these are culled from each of the stories. Little by little, categories are shaped by the content of the stories themselves. In this way the respondent furnishes the dimensions of analysis as well as the scale values on these dimensions.

Second Test. It is possible to wonder whether it is true that the opprobrium that is heaped on the Nescafé woman comes from her use of a device that represents a short cut and labor-saver in an area where she is expected to embrace painstaking time-consuming work in a ritualistic way. To test this a variation was introduced into the shopping

TABLE 1
PERSONALITY CHARACTERISTICS ASCRIBED TO USERS OF PREPARED FOODS

<i>If They Use</i>	<i>No Prepared Food (Maxwell House Alone)</i>		<i>Nescafé (Alone)</i>		<i>Maxwell House (Plus Pie Mix)</i>		<i>Nescafé (Plus Pie Mix)</i>	
	<i>Number</i>	<i>Per Cent</i>	<i>Number</i>	<i>Per Cent</i>	<i>Number</i>	<i>Per Cent</i>	<i>Number</i>	<i>Per Cent</i>
They are seen as:								
Not Economical	12	17	24	32	6	30	7	35
Lazy	8	11	46	62	5	25	8	40
Poor Personality and Appearance	28	39	39	53	7	35	8	40
N =	72		74		20		20	

lists. In a second experiment, 150 housewives were tested with the form given above, but a sample was added to this group which responded to a slightly different form. If we assume that the rejection in the first experiment came from the presence of a feeling about synthetic shortcuts, we might assume also that the addition of one more shortcut to both lists would bring the Maxwell House woman more into line with the Nescafé woman, since the former would now have the same guilt that the Nescafé woman originally had, while the Nescafé woman, already convicted of evading her duties, would be little further injured.

In order to accomplish this, a second prepared food was added to both lists. Immediately after the coffee in both lists the fictitious

item "Blueberry Fill Pie Mix" was added. The results are shown in Table 1.

It will be seen immediately, in the first two columns, that the group to whom the original form of the list was given showed the same kind of difference as reported above in their estimates of the two women. The group with an additional prepared food, however, brought the Maxwell House Coffee woman down until she is virtually undistinguishable from the Nescafé. There seems to be little doubt but that the prepared-food-character, and the stigma of avoiding housewifely duties is responsible for the projected personality characteristics.

Relation to Purchasing. It is still relevant to ask whether the existence of these feelings in a potential consumer is related to purchasing. It is hypothesized that these personality descriptions provide an op-

TABLE 2

<i>The woman who buys Nescafé is seen as:</i>	<i>By Women Who Had Instant Coffee in the House (N = 32)</i>		<i>By Women Who Did Not Have Instant Coffee in the House (N = 18)</i>	
	<i>Number</i>	<i>Per Cent</i>	<i>Number</i>	<i>Per Cent</i>
Economical**	22	70	5	28
Not economical	0	0	2	11
Can not cook or does not like to**	5	16	10	55
Plans balanced meals*	9	29	2	11
Good housewife, plans well, cares about family**	9	29	0	0
Poor housewife, does not plan well, does not care about family*	5	16	7	39
Lazy*	6	19	7	39

* A single asterisk indicates that differences this great would be observed only 5 times out of 100 in repeated samplings of a population whose true difference is zero.

** A double asterisk indicates that the chances are 1 in 100. We are justified in rejecting the hypothesis that there is no difference between the groups.

portunity for the consumer to project hopes and fears and anxieties that are relevant to the way the product is seen, and that they represent important parts of her motivation in buying or not buying. To test this hypothesis, a small sample of fifty housewives, comparable in every way to the group just referred to, was given the original form of the shopping list (Nescafé only). In addition to obtaining the personality description, the interviewer, on a pretext, obtained permission to look at her pantry shelves and determine personally whether or not she had instant coffee of any brand. The results of this investigation are shown in Table 2.

The trend of these data shows conclusively that if a respondent sees the woman who buys Nescafé as having undesirable traits, she is not likely to buy instant coffee herself. The projected unacceptable

characteristics go with failure to buy, and it does not seem unwarranted to assume that the association is causal.

Furthermore, these projected traits are, to some extent, additive. For instance, if a respondent describes the woman as having one bad trait only, she is about twice as likely not to have instant coffee. However, if she sees her as having two bad traits, and no good ones (for example, lazy, cannot cook), she is about three times as likely not to have instant coffee as she is to have it. On the other hand, if she sees her as having two good traits (such as, economical, cares for family), she is about six times as likely to have it as not.

It was pointed out earlier that some women felt it necessary to "excuse" the woman who bought Nescafé by suggesting that she lived alone and hence could not be expected to be interested in cooking, or that she had a job and did not have time to shop better. Women who had instant coffee in the house found excuses almost twice as often as those who did not use instant coffee (12 out of 32, or 42 per cent, against 4 out of 18, or 22 per cent). These "excuses" are vitally important for merchandising. The need for an excuse shows that there is a barrier to buying in the consumer's mind. The presence of excuses shows that there is a way around the barrier. The content of the excuses themselves provides valuable clues for directing appeals toward reducing buying resistance.

CONCLUSIONS

There seems to be no question that in the experimental situation described here: (1) motives exist which are below the level of verbalization because they are socially unacceptable, difficult to verbalize cogently, or unrecognized; (2) these motives are intimately related to the decision to purchase or not to purchase; and (3) it is possible to identify and assess such motives by approaching them indirectly.

Two important general points come out of the work reported. The first is in the statement of the problem. It is necessary for us to see a product in terms of a set of characteristics and attributes which are part of the consumer's "private world," and as such may have no simple relationship to characteristics of the object in the "real" world. Each of us lives in a world which is composed of more than physical things and people. It is made up of goals, paths to goals, barriers, threats, and the like, and an individual's behavior is oriented with respect to these characteristics as much as to the "objective" ones. In the area of merchandising, a product's character of being seen as a path to a goal is usually very much more important as a determinant of purchasing than its physical dimensions. We have taken advantage

of these qualities in advertising and merchandising for a long time by an intuitive sort of "playing-by-ear" on the subjective aspects of products. It is time for a systematic attack on the problem of the phenomenological description of objects. What kinds of dimensions are relevant to this world of goals and paths and barriers? What kind of terms will fit the phenomenological characteristics of an object in the same sense that the centimeter-gram-second system fits its physical dimensions? We need to know the answers to such questions, and the psychological definitions of valued objects.

The second general point is the methodological one that it is possible, by using appropriate techniques, to find out from the respondent what the phenomenological characteristics of various objects may be. By and large, a direct approach to this problem in terms of straightforward questions will not yield satisfactory answers. It is possible, however, by the use of indirect techniques, to get the consumer to provide, quite unselfconsciously, a description of the value-character of objects in his environment.

11. PROJECTIVE TECHNIQUES IN ATTITUDE MEASUREMENT*

Though concerned primarily with sociological and public opinion issues, the applications of projective techniques reviewed in this article are sufficiently close to commercial research situations to suggest numerous uses in this area as well. In this sense, the large number and variety of studies reviewed, as indicated in part by the extensive references in the footnotes, considerably enhance the value of this article. The techniques are not presented in any great detail, which is understandable in this type of article. Sufficient descriptive material is provided in each case, however, that the specific approach used becomes quite clear, especially with the preceding article as a guide.

A few years ago McNemar published a monograph¹ in which he stated that opinion or attitude techniques were a simple matter of

* Adapted from an article by Irving R. Weschler, Institute of Industrial Relations, University of California, Los Angeles, and Raymond E. Bernberg, Los Angeles State College, "Indirect Methods of Attitude Measurement," *International Journal of Opinion and Attitude Research*, Vol. 4 (Summer, 1950), pp. 209-28. See also Irving R. Weschler, "Problems in the Use of Indirect Methods of Attitude Measurements," *Public Opinion Quarterly*, Vol. 15 (Spring, 1951), pp. 133-38.

¹ Q. McNemar, "Opinion-Attitude Methodology," *Psychological Bulletin*, Vol. 43 (July, 1946), pp. 289-374.

asking people questions about an issue in order to elicit a response which can be interpreted as the respondent's attitude toward the given issue. "The introduction of high sounding definitions," wrote McNemar, "does not alter the fundamental fact that practically all attempts at determining opinion are at the verbal level—the correlation between verbal and overt non-verbal behavior, the latter depending in part upon the opinion or attitudinal set, is an unknown and is usually left as an unknown by investigators in this field."

The article deplored the lack of a sound scientific methodology and repeatedly stressed the need for reliability, validity, and unidimensionality for the instruments or devices used to classify or measure people with respect to their opinions or attitudes.

McNemar's description of the field of attitude measurement seems to be unduly narrow and restricted, since an attitude as a psychological construct may be observed in both verbal and nonverbal behavior. It exists only as part of a whole situation and requires forces external to the individual which demand response. An attitude may be considered as the psychological counterpart of the dynamic processes within the individual which result in the organization of his perception of any given situation and are manifested in his response. Thus, a raised eyebrow or a shrug of the shoulder are as indicative of an attitude as a verbal response which is intended to express an opinion.

Historically, two opposing points of view have been adopted by writers and research workers with regard to the range and depth of attitudes. Bogardus,² Symonds,³ Hartshorne,⁴ and many others maintain that attitudes are *specific* and as numerous as the objects to which persons respond. Other authors, such as Faris,⁵ Dewey,⁶ and Likert⁷ consider that attitudes represent a number of isolated dispositions and are therefore *general* in nature. Both sides have adduced experimental evidence to support their view, but in present-day research the possible dichotomy between specific and general attitudes is no longer the major focus of interest. Modern social psychology concerns itself

² E. S. Bogardus, *Fundamentals of Social Psychology* (New York: Appleton Century, 1931).

³ P. M. Symonds, "What Is an Attitude?" *Psychological Bulletin*, Vol. 24 (March, 1927), pp. 200-201.

⁴ H. Hartshorne, M. A. May, and F. K. Shuttlesworth, *Studies in the Nature of Character* (New York: Macmillan Company, 1930).

⁵ G. Faris, "The Concept of Social Attitudes," in K. Young (ed.), *Social Attitudes* (New York: Henry Holt & Co., 1931), pp. 3-16.

⁶ J. Dewey, *Human Nature and Conduct* (New York: Henry Holt & Company, 1948).

⁷ R. Likert, "A Technique for the Measurement of Attitudes," *Archives of Psychology*, Vol. 22 (June, 1932), No. 140.

mainly with the role of behavior as a function of attitudes, and the emphasis is not on their make-up but rather on their potential effects.

The term "attitude" implies essentially a disposition to respond as a function of the whole situation. In view of this orientation, it is the stimulus or forces demanding response which may be considered specific or general in nature rather than the attitudes which they provoke, and the attitudes themselves may be held to be predictable when the conditions arousing them are known. The role of attitude measurement is therefore to utilize appropriate sampling techniques and experimental controls in order to establish relationships with these impinging stimuli and with the dimensions of response within which attitudes may be logically set as variables.

A great deal of criticism has been directed against the use of certain attitude measurement techniques, especially against those based on simple scales and on direct questioning, because they deal only with manifest verbal content and fail to reach into the more comprehensive aspects of the personality. The person who is asked point-blank to express his feelings on a subject about which he is reticent for one reason or another many well evade the issue by providing an answer which conforms with the views of the investigator or which is sufficiently neutral to protect his psychological security. This process does not have to be conscious or intentional, and many clinical studies have shown that certain attitudes, although no less real to the individual, have been suppressed for being unacceptable to his values and standards and become inaccessible to the explicit frontal approach of the various direct measurement techniques.

It is for the purpose of exploring these "deeper levels" that other attitude measurement devices are needed which enable the experimenter to "observe the manifestations of the attitude as it affects action and perception and judgment of other objects which are somehow functionally connected with the object in question."⁸ The advantage of such an "indirect" method of measurement lies in the fact that it conceals from the individual the intent of the measurement and allows him to produce his responses freely without fear of getting personally involved. Furthermore, such indirect measurement enables the experimenter to observe and measure without producing an effect on the attitude itself; in no way is the situation structured so as to force the subject consciously to reveal his bias, and the measurement consists essentially of the quantitative interpretation of responses which are

⁸ D. Krech and R. S. Crutchfield, *Theory and Problems of Social Psychology* (New York: McGraw-Hill Book Co., 1948), p. 246.

considered valuable for shedding light upon the underlying attitudes.

It is important to remember that, although the subject is to be unhampered in his performance, the identification and the measurement of attitudes require that his responses be analyzed both in terms of their qualitative nature and quantitative data.

The criteria which McNemar considered essential for good scaling on questionnaire methods, such as reliability and validity, do apply equally well to the various "indirect" techniques which will be described in the following pages. In their present stage of development, these "indirect" methods have not yet attained the wide range of applicability of their forerunners, but it would seem that the further use and perfection of these tools provide an extremely fruitful area for future methodological research.

Among the "indirect" methods which lend themselves especially well to the measuring of social attitudes are certain projective techniques which allow the subject free scope to action and provide him with the widest possible latitude in choice of response or form of expression. A variety of media have been used. These range from the application of a completely unstructured stimulus, such as an ink blot, to those which are highly structured, such as selected photographs or cartoons.

In an evaluation of the technology of attitude measurement, the most fruitful classification of "indirect" methods seems to be in terms of the material presented to the subject, because it facilitates the study of the revelation of attitudes as they are brought into the open by the application of many stimulus devices. The classification which is suggested is by no means complete but represents a flexible and useful grouping according to the nature of the materials used. In essence, these materials can be classified as "pictorial," "play and dramatic," and "verbal and written."

PICTORIAL MATERIALS

In a pioneering study, Horowitz undertook a series of experiments which provided three indirect testing devices to reveal the attitude of a subject.⁹ The first was the "Ranks" test containing pictures of boys' faces, white and Negro, for the purpose of racial judgments; the second was the "Show Me" test, consisting of the same pictures, but requiring a different judgment; the third was the "Social Situations" test con-

⁹ R. E. Horowitz, "Racial Aspects of Self-Identification in Nursery School Children," *Journal of Psychology*, Vol. 7 (1939), pp. 91-99.

sisting of 15 posed situations with white boys in the pictures at times, and Negro boys taking their places in other pictures. This is similar to the earlier Horowitz study using pictures to get at the attitudes of white children toward the Negro.¹⁰

Seeleman utilized the technique of pictorial materials in the effects of attitudes upon memory.¹¹ She used pictures of Negroes and whites in conjunction with a Negro attitude scale to determine differences in memory of picture recognition related to favorable and unfavorable attitudes toward the Negro. The unfavorable attitude tended to obliterate recognition of individual differences among Negro pictures, whereas a favorable attitude tended to heighten recognition of these differences.

The most systematic analysis of the content of pictorial materials was undertaken by Murray and his co-workers in a study emphasizing the clinical approach to the development of attitudes and beliefs.¹² These investigations describe the use of a variety of projective techniques, stressing especially the development of the Thematic Apperception Test. Similar clinical techniques were used by Frenkel-Brunswik and Sanford in their study of the "anti-Semitic personality,"¹³ and by French in her thorough investigation of the personality pattern of 15 college girls.¹⁴

Murray and Morgan utilized a variety of new indirect measurement devices in their clinical study of sentiments toward war, religion, parents, and sex.¹⁵ In designing their methodology, the course followed was that of listing the more important common "foci of sentiment," up to 100 or so, and then introducing as many of them as possible into each test so that the principal sentiments could be tapped by several different techniques. The instruments which made use of pictorial materials included the so-called "Picture Selection Test"

¹⁰ E. L. Horowitz, "The Development of Attitude toward the Negro," *Archives of Psychology*, No. 194 (January, 1936).

¹¹ V. Seeleman, "The Influence of Attitude upon the Remembering of Pictorial Material," *Archives of Psychology*, No. 258 (September, 1940).

¹² H. A. Murray, *et al.*, *Explorations in Personality* (New York: Oxford University Press, 1938).

¹³ E. Frenkel-Brunswik and R. N. Sanford, "Some Personality Factors in Anti-Semitism," *Journal of Psychology*, Vol. 20 (1945), pp. 271-91.

¹⁴ V. V. French, "The Structure of Sentiments," *Journal of Personality*, Vol. 15 (June, 1947), pp. 247-82; Vol. 16 (September, 1947), pp. 78-108.

¹⁵ H. A. Murray and C. D. Morgan, "A Clinical Study of Sentiments," *Genetic Psychology Monographs*, No. I (August, 1945), pp. 1-149; No. II (November, 1945), pp. 153-309.

which consisted of 225 pictures cut out of various illustrated magazines, depicting 45 "foci of sentiments," such as Old Man, "Fast" Woman, Children, Laborers, Negroes, and many others. Each "focus" was represented by eight pictures, and more than half of the illustrations portrayed two or more different "foci."

Murray and Morgan based the utility of this test upon the empirically determined fact that, in judging pictures of nearly equal aesthetic worth, people will be guided largely by the appeal or repulsion of the subject matter. If a subject repeatedly selects pictures belonging to the same category, it seems highly probable that a strong sentiment is operating.

Proshansky investigated attitudes toward organized labor by means of a modified version of the Thematic Apperception Test.¹⁶ His hypothesis was that extreme groups, that is, those having strong pro-labor or anti-labor attitudes, would reveal their social orientation through their manner of report upon pictures of social situations. Using a series of pictures which he took from periodicals current at the time, the author selected a set which, in the judgment of three judges, were ambiguous with respect to outcome as far as labor was concerned, that is, indicating neither victory nor defeat for the labor cause. A few other pictures which did not deal with industrial situations were added to disguise the purpose of the test.

The method, according to Proshansky, clearly permitted outside distortion at the time of the original perception, or retrospective falsification as attitude got in its work upon memory, or sheer elaboration of the meaning of the picture, consciously going beyond anything that the picture actually offered. The responses were evaluated, in regard to their expressed attitude toward labor, by three trained judges and high correlations were obtained between these analyzed picture responses and the score on the attitude scale.

Loebowitz-Lennard and Riessman have recently reported the development of a "social perception" test.¹⁷ This test consists of a series of pictures dealing with all areas of social interaction, which emphasizes particularly the dynamics of Negro-white and Jew-Gentile inter-group relations. Similarly, Sollenberger and Pulford have made a comparative study of white and Negro children, using the Thematic

¹⁶ H. M. Proshansky, "A Projective Method for the Study of Attitudes," *Journal of Abnormal and Social Psychology*, Vol. 38 (July, 1943), pp. 393-95.

¹⁷ H. Loebowitz-Lennard and F. Riessman, Jr., "A Preliminary Report on a Social Perception Test: A New Approach to Attitude Research," *Social Forces*, Vol. 24 (May, 1946), pp. 423-27.

Apperception Test with concomitant questionnaires and interviews to get at the learning process of racial prejudices.¹⁸

Another useful application of pictorial materials is the "cartoon technique," which was developed in varying forms by Rosenzweig¹⁹ and by Fromme.²⁰ The latter utilized it as part of his intensive study of opinions on preventing war. The materials consisted of a series of five political cartoons from which the original captions were removed. Four other titles, expressing a famous quotation or proverb, or describing the picture with a "catch-all" word, were supplied beneath each picture, and the subject's task was to choose what he thought to be the most appropriate caption. The investigation showed that attitudes could be revealed in this manner and the discussion which was inevitably provoked by these pictures helped to clarify the subject's position. There was a slight tendency evident for subjects addicted somewhat more extremely than others to a particular viewpoint to react to the meaning of the caption more specifically than to its appropriateness for the cartoon. The justification for the subject's choice, although usually stated in terms of the material provided, gave a good indication of the attitudinal forces at work.

A modification of the Rosenzweig technique was utilized by J. F. Brown to cast light on the presence of racial hostile attitudes.²¹ In this case, the material of the test, like its Rosenzweig prototype, consisted of a series of 24 cartoon-like pictures, each depicting a frustrating situation of common occurrence. The left of each picture showed a "frustrating" person saying certain words, which either helped to describe the frustration in which another individual on the right of the picture found himself, or which of themselves actually frustrated this other person. Facial and other expressions of personality were purposely left out. The person on the right was always shown with a blank caption box above, which was used by the subject to record his "projected" response.

This test aimed particularly at studying aggressive reactions of in-

¹⁸ R. T. Sollenberger and E. H. Pulford, "Attitudes of White Children toward Negroes in Social Situations as Measured by a Thematic Apperception Technique," *Bulletin 2, Inventory of Research in Racial and Cultural Relations*, American Council on Race Relations, Vol. 1 (1948), pp. 61-62.

¹⁹ S. Rosenzweig, "The Picture-Association Method and Its Application in a Study of Reactions to Frustration," *Journal of Personality*, Vol. 14 (September, 1945), pp. 3-23.

²⁰ A. Fromme, "On the Use of Certain Qualitative Methods of Attitude Research: A Study of Opinions on the Methods of Preventing War," *Journal of Social Psychology*, Vol. 13 (May, 1941), pp. 429-59.

²¹ J. F. Brown, "A Modification of the Rosenzweig Picture-Frustration Test to Study Hostile Interracial Attitudes," *Journal of Psychology*, Vol. 24 (July, 1947), pp. 247-72.

dividuals to frustrating situations in which the person frustrated or frustrating is obviously a member of one of the minority groups, that is, is clearly characterized as Jew or Negro. Although the results reported by Brown were based on a small number of subjects, an analysis of the protocols of various individuals with known attitudes pointed to this method's considerable usefulness for the study of inter-racial hostility.

PLAY AND DRAMATIC MATERIALS

Projective techniques represent a method of investigating the personality organization of an individual and facilitate a comprehensive understanding of human expression. The projective response is always a part of the person who produced it, but the exact pattern of circumstances which was responsible for its utterance is never known. Symonds deduces that projective techniques in their modern use cannot be thought of as tests which tell something definite and clear about a person.²² He feels, rather, that they suggest hypotheses which must be substantiated by referring them to actual life history or interview material or to the subject's performance on statistically reliable objective instruments. A number of methods for the validation of projective techniques have been proposed by Macfarlane,²³ but none of these has as yet produced the kind of data which lend themselves to an adequate quantitative treatment.

Among the techniques which were found to elicit social attitudes by "indirect" means, the use of play and dramatic materials has offered interesting possibilities. The majority of these investigations allowed the subjects to participate actively in the stimulus situation, with the interpretation of their behavior greatly dependent upon the skill of the experimenter. As yet, none of the materials have been sufficiently standardized so that they can be economically and easily applied by technicians in everyday nonlaboratory situations. But even though the imagination of the researchers has advanced faster than their technology, important strides forward have been made in the small-scale clinical and laboratory measurement of attitudes.

Bell, who has published a careful analysis of projective techniques,²⁴ cites a series of studies by Baruch, who in semistructured play interviews with 46 children gained data on their attitudes, especially of

²² P. M. Symonds, "Projective Techniques," in P. L. Harriman, *Encyclopedia of Psychology* (New York: Philosophical Library, 1946), pp. 583-90.

²³ J. W. Macfarlane, "Problems of Validation Inherent in Projective Methods," *American Journal of Orthopsychiatry*, Vol. 12 (July, 1942), pp. 405-12.

²⁴ J. E. Bell, *Projective Techniques* (New York: Longmans-Green, 1948).

aggression to their respective families.²⁵ Her equipment was simple—a box of dolls representing different ages of both sexes, a bed, an armchair, a toilet, and a piece of India print the size of a handkerchief to make out a room. Dolls duplicating the family constellation of each child were chosen, and a 15-minute interview was held in which the child was told he might do anything he likes with them, or on another occasion, that he could be as mean as he cared. The results indicated the presence of aggression in two-thirds of the children, found at all ages from two to five years. The children revealed this aggression by separating certain members of the family constellation from others, spanking, burying, crushing, and twisting the dolls, calling them names, and drowning them in the toilet.

In a manner similar to E. L. Horowitz, Clark and Clark investigated the genesis of racial attitudes and ego-identification in Negro children.²⁶ Among the techniques which they employed was a modification of the Horowitz Show-Me test, using dolls instead of pictures.

Homburger used a variation of the doll technique with adults.²⁷ He presented his subjects (college freshmen) with a group of dolls, household furniture, blocks, and cars and asked them to prepare a "dramatic scene." The scene which the subject arranged was recorded photographically and analyzed in terms of psychoanalytic symbols apparent in the arrangement of the material.

More recently, Evans and Chein have devised a "Movie Story game" for the projective testing of interracial attitudes of Negro and white children.²⁸ The children are not aware of the true nature of the test. Two miniature stages and a group of miniature figures, half Negro and half white, are used. One figure is selected as the main character and is of the same color, sex, and age as the subject; the subject is asked to tell what the main character would do or say in various standard situations. The test is followed by a group of social distance type questions and an interview. This is still in the preliminary stage of development.

²⁵ D. W. Baruch, "Aggression during Doll Play in a Pre-School," *American Journal of Orthopsychiatry*, Vol. 11 (April, 1941), pp. 252-59.

²⁶ K. B. Clark and M. P. Clark, "Racial Identification and Preference in Negro Children," in T. M. Newcomb and E. L. Hartley (eds.), *Readings in Social Psychology* (New York: Henry Holt & Co., 1947), pp. 169-78.

²⁷ E. Homburger, "Section on 'Play Therapy.'" *American Journal of Orthopsychiatry*, Vol. 8 (July, 1938), pp. 507-10.

²⁸ M. C. Evans and I. Chein, "The Movie Story Game: A Projective Test of Interracial Attitudes for Use with Negro and White Children," Bulletin 2, *Inventory of Research in Racial and Cultural Relations*, American Council on Race Relations (1948), p. 6.

Some other studies have dealt with the problem of social adjustment and the analysis of attitudes on the basis of play behavior. Typical among these investigations are the work of Anderson,²⁹ Hanfmann,³⁰ Horowitz and Horowitz,³¹ and Wright.³²

The play situation which for purposes of attitude measurement has the highest degree of standardization is the so-called "World test," developed by Lowenfeld³³ and Buehler and Kelly,³⁴ and systematically scored by Bolgar and Fischer.³⁵ In essence the equipment consists of a box with miniature wooden toys representing houses, trees, fences, cars, people, animals, and so on, and the subject is asked to construct whatever he would like on a wooden or sand tray. The validity of the test was checked against complete biographies and found to be high. The scoring scheme is divided into six categories and considers such factors as the order of choice, the quantity of pieces selected, the nature of the choice as indicated by "the Gestalt achieved," the contents of the items used, the behavior of the subjects, and finally the verbalization of the subjects which took place during the testing period.

In a clinical play study which was especially structured to elicit social attitude, Dubin found that it was possible to make fairly good predictions of scores on attitude scales toward labor, government, or war on the basis of the individual's play construction with a set of toys.³⁶ The subjects, six women students of Barnard College and four men students of Columbia College, were provided with 80 toys representing four functional groups—war, public service, labor, and travel and entertainment—and were instructed to represent the world as they saw it and the world as they would like to see it. An interview

²⁹ H. H. Anderson, "Domination and Integration in the Social Behavior of Young Children in an Experimental Play Situation," *Genetic Psychology Monographs*, Vol. 19 (1937), pp. 341-408.

³⁰ E. Hanfmann, "Social Structure of a Group of Kindergarten Children," *American Journal of Orthopsychiatry*, Vol. 5 (October, 1935), pp. 407-10.

³¹ E. L. Horowitz and R. E. Horowitz, "Development of Social Attitudes in Children," *Sociometry*, Vol. 1 (January, 1938), pp. 301-38.

³² B. A. Wright, "An Experimentally Created Conflict Expressed by Means of a Projective Technique," *Journal of Social Psychology*, Vol. 21 (May, 1945), pp. 229-45.

³³ M. Lowenfeld, "The Theory and Use of Play in the Psychotherapy of Childhood," *Journal of Mental Science*, Vol. 84 (November, 1938), pp. 1057-58.

³⁴ C. Buehler and G. Kelly (eds.), *The World Test* (New York: Psychological Corporation, 1941).

³⁵ H. Bolgar and L. Fischer, "The Toy Test: A Psychodiagnostic Method," *Psychological Bulletin*, Vol. 37 (July, 1940), pp. 517-18 (abstract).

³⁶ S. S. Dubin, "Verbal Attitudes Scores from Responses Obtained in a Projective Technique," *Sociometry*, Vol. 3 (January, 1940), pp. 24-48.

was held with each subject where he interpreted the worlds he made. Three graduate students in psychology were given pictures of "the worlds" and asked to predict on a 5-point scale the responses of the students on "A Survey of Opinions," modeled after Murphy and Likert.³⁷ The comparison between the scores predicted for the subjects on the attitude scales and the scores actually made by them resulted in a mean coefficient of correlation of .49, which seems highly satisfactory considering the nature of the experiment and invites further research with larger groups and with different opinion scales.

The subject's participation in the play situation reaches its broadest characterization in the various "role playing" techniques developed by Moreno and his associates.³⁸ These dramatic techniques, standardized in method under the heading of "psychodrama," were originally developed for purposes of individual diagnosis and therapy in psychiatric cases but today are the foundation of a variety of methodologically different investigations and the backdrop for numerous "theoretical formulations mixed with mystic notions."³⁹

The use of the psychodramatic test has been summarized by Bell.⁴⁰ It consists of placing the individual on stage after a brief interview with the director and observing and evaluating his performance in a number of artificially created enterprises. Within the test situations, a wide range of themes can be introduced by the director, including love, death, economic problems, status, security, and social attitudes. The variety of themes and roles permits the director to "reach for the limits" in the subject's behavior. The psychodramatic interview concludes with the analysis of the responses in terms of their imaginal content, their perceptual characteristics, their degree of involvement and organization, and finally their manifest social interaction.

Discussion of Moreno's contributions would not be complete without mention of the valuable techniques which he developed for the portrayal and measurement of the complex interrelationships that exist even among the very smallest groups. Sociometry⁴¹ consists of a set of objective techniques whose aim it is to establish the pattern of feelings of acceptance and rejection or like and dislike that exist among the

³⁷ G. Murphy and R. Likert, *Public Opinion and the Individual* (New York: Harper and Brothers, 1938).

³⁸ J. L. Moreno, *Psychodrama*, Vol. 1 (New York: Beacon House, 1946).

³⁹ M. Sherif and H. Cantril, *The Psychology of Ego-Involvements* (New York: John Wiley & Sons, 1947), p. 341.

⁴⁰ Bell, *op. cit.*

⁴¹ J. L. Moreno, *Sociodrama: A Method for the Analysis of Social Conflicts* (New York: Beacon House, 1944).

members of the group. The thesis is that the most significant social groupings are those based on such feelings rather than those based on formal structures. Moreno believes that an analysis of the external frame of the group fails to reveal the "spontaneous" groupings which are psychologically more real than some of their formal counterparts. Although the methodology of measuring personal attitudes is direct and based upon the expressed likes and dislikes of the individual members in the group for each other, the constellation of the total attitudes of the group itself is arrived at through ingenious formulations of a more or less indirect nature.

VERBAL AND WRITTEN MATERIALS

The value of an indirect approach to attitude measurement results from the fact that an individual's overt behavior frequently does not correspond to his stated attitude. Thus, on a public opinion poll, a person may respond in one way and then act in quite another. It is further believed that the more controversial an issue becomes, the less likely is the average individual to express his beliefs freely and candidly.

In an effort to study the frankness of his subjects, Spencer administered a personality tool, known as the Experience Appraisal Blank,⁴² to a group of 192 high school students who were instructed not to sign their names and were assured that their responses would be kept confidential because their identities would not be known. A later inquiry revealed that 22 per cent of the subjects would have left some of the items unanswered if their signatures had been requested, and about 9 per cent admitted that they would have answered the questions untruthfully.

On the other hand, a study by Robinson and Rohde, which dealt specifically with a comparison of the effectiveness of direct and indirect methodology in the measurement of attitudes, produced different results.⁴³ They found that anti-Semitic responses were more numerous in answer to direct questions than to indirect questions. The possibility was suggested that anti-Semitism takes a somewhat latent form, in stereotyped opinions, and that these are not produced by subtle or indirect questions. There seemed to have been less readiness to attack the Jews on any occasion than there was to agree with anti-Jewish

⁴² D. Spencer, "The Frankness of Subjects on Personality Measures," *Journal of Educational Psychology*, Vol. 29 (January, 1938), pp. 26-35.

⁴³ D. Robinson and S. Rohde, "Two Experiments with an Anti-Semitism Poll," *Journal of Abnormal and Social Psychology*, Vol. 41 (April, 1946), pp. 136-45.

statements when they were explicitly stated. This hypothesis seems reasonable, provided it is recognized that the total situation presented no threat to the psychological security of the subjects who were questioned and whose degree of ego-involvement was of a relatively minor character.

In a series of experiments, Morgan and Morton succeeded in showing that the reasoning process of an individual may be modified by underlying viewpoints which may or may not agree with the opinions he expresses overtly.⁴⁴ These distortions of reasoning can then be utilized as an indirect index of the underlying attitudes of the respondents. The method consists of a comparison of answers to a syllogism of neutral emotional value with those answers to a similar syllogism involving a controversial issue. The test is set up in such a manner that the responses of the subjects on the controversial syllogisms can be compared to reactions on the syllogisms of a neutral nature which serve as controls.

The role of verbal distortion in the transmission of rumors was studied by Allport and Postman in another series of highly dramatic experiments.⁴⁵ With the aid of a set of potentially exciting and ambiguous stimulus pictures, these investigators were able to observe and describe the factors which operate through time in the verbal organization and restructuring of their subjects' perception.

Their procedure consisted, in essence, of the following: A slide depicting some detailed situation was thrown on the screen, and one subject was asked to describe it, while looking at it, to another subject, who was unable to see the stimulus. This second individual in turn had to rely on his memory to relate verbally the description to a third subject, and so on down the line. The last person was then asked to repeat the story which he had heard, and his version of the picture was critically compared with the original on the screen.

Allport and Postman interpreted their findings mainly in terms of the changes which occurred in the verbal description during the process of transmission, dealing incidentally with the motivational and attitudinal components which were responsible for the subjects' structuring of the materials. The evidence from other experiments in this area points to the possibility that the distortions in the transmis-

⁴⁴ J. J. B. Morgan and J. T. Morton, "Distorted Reasoning as an Index of Public Opinion," *School and Society*, Vol. 57 (March 20, 1943), pp. 333-35; J. J. B. Morgan and J. T. Morton, "The Distortion of Syllogistic Reasoning Produced by Personal Convictions," *Journal of Social Psychology*, Vol. 20 (August, 1944), pp. 39-59.

⁴⁵ G. W. Allport and L. Postman, *The Psychology of Rumor* (New York: Henry Holt & Co., 1947).

sion process are greatly influenced by the attitudes operating in the situations and that the key to those attitudes can be found in an examination of the subjects' expressed and repressed verbal behavior. It seems plausible that a step-by-step analysis of the protocols can be made which would allow the trained investigator to analyze the responses in terms of the attitudinal contributions or omissions in each phase of the transmission chain and to categorize them with the aid of certain modified Thematic Apperception Test scoring techniques.

The relationship between belief in rumors and expressed attitudes was systematically explored by Smith.⁴⁶ In an experiment to determine the influence on belief in unconfirmed news-type statements about Russia by labeling them "facts" and "rumors," he found that the attitude of the subject as measured by a regular attitude scale was reflected in his belief scores. Thus the pro-Soviet people placed the greatest credence in the pro-Soviet news statements, under all conditions of labeling, while the pronounced anti-Soviet people exhibited the least acceptance of these statements. The investigator reached the conclusion that the absolute amount of belief an individual places in the "news items" under any condition of labeling almost certainly depends in part on the initial attitude with which he approaches the items. The findings of this experiment again open the possibility of quantitatively measuring a person's attitudes by establishing a ratio between the number of accepted and rejected controversial statements which may be contained in a test of "rumor proneness."

A similar technique which would utilize the mechanism of projection and the universal tendency toward autistic selection and recall of information was proposed by Loebowitz-Lennard and Riessman.⁴⁷ These investigators planned to provide their subjects with a series of unstructured questions on social issues the answers to which would require them to project their existing attitudes.

Of all the verbal devices which have been utilized to elicit attitudes in a more or less indirect manner, the word association technique is probably the oldest. In its simplest form it consists of presenting a list of words to the subject and asking him to respond with the first word that comes to mind. Bell gives a detailed description of

⁴⁶ G. H. Smith, "Beliefs in Statements Labeled Fact and Rumor," *Journal of Abnormal and Social Psychology*, Vol. 42 (January, 1947), pp. 80-90; "The Relation of 'Enlightenment' to Liberal-Conservative Opinions," *Journal of Social Psychology*, Vol. 28 (August, 1948), pp. 3-17.

⁴⁷ H. Loebowitz-Lennard and F. Riessman, Jr., "A Proposed Projective Attitude Test," *Psychiatry*, Vol. 9 (February, 1946), pp. 67-68.

the variety of purposes for which this technique has been employed.⁴⁸ Recently, the word association technique was used in a modified form by Murray and Morgan in their study of sentiments toward war, religion, parents, and sex.⁴⁹ As part of their Sentiments Examination, the examiner is required to read off a list of 48 words (such as Father, Communism, Religion), to each of which the subject is asked to respond by giving the most descriptive adjectives he can think of. The subject is led to believe that the examiner is interested in testing the range of his vocabulary, but the attitudinal analysis is carried out by determining the ratio of appreciative to depreciative adjectives.

An alternative method to word association is the sentence completion test, which can present a more personalized medium for the projection of significant themes. It also provides more material for analysis since the responses are longer and possibly more meaningful than those obtained through word association.

The sentence completion method was first developed by Tendler,⁵⁰ who believed that the presentation of stimuli in the form of incomplete sentences would arouse a particular emotional set and yet allow for free responses. In analyzing the responses of 250 college girls, he noted that the same stimulus evokes different responses in different individuals, that individuals differ in the associative flow of responses, and that the responses provide a key to the subject's fears, likes, interests, attachments, and attitudes.

A sentence completion test which can be changed to meet the specific situation was constructed for use in military psychiatric centers by Shor.⁵¹ The stimuli which he utilized were the beginnings of grammatical sentences suggesting contents, feeling tones, qualities of attitude, and specific objects as areas of attention. These stimuli were arranged in such a sequence as to permit a carry-over or generalization of attitude from immediate to basic human interests. Fifteen per cent of the items reflected the military situation, thirty-five per cent dealt with attitudes frequently experienced in relation to military life and yet normally associated with civilian life, and the rest covered general life situations.

A freer reign of expression is provided through a series of projec-

⁴⁸ Bell, *op. cit.*

⁴⁹ Murray and Morgan, *op. cit.*

⁵⁰ A. D. Tendler, "A Preliminary Report on a Test for Emotional Insight," *Journal of Applied Psychology*, Vol. 14 (April, 1930), pp. 122-36.

⁵¹ J. Shor, "Report on a Verbal Projective Technique," *Journal of Clinical Psychology*, Vol. 2 (July, 1946), pp. 279-82.

tive methods which allow the subject to call upon his literary imagination and to project his feelings and attitudes into the completion of arguments, the invention of story plots, or the writing of autobiographical materials. The common element in the techniques suggested is the provision of a theme for elaboration, with the principles of interpretation and scoring being similar to those of the Thematic Apperception or other pictorial tests.

Typical of these devices is the Argument Completion, developed by Murray and Morgan.⁵² The subject is given a card, on which is printed a brief description of two young men (X and Y) who are engaged in an argument, and he is asked to continue and finish the argument in the form of a realistic dialogue. The unsolved arguments are presented on ten separate cards, each dealing with a different sentiment. The subject is usually led to believe that his powers of argumentation are being tested and becomes quickly involved in the controversy which he is inventing. Murray and Morgan found the technique useful because it usually forces the subject to explore more of his own sentiments than he might otherwise be inclined to do. In most cases it is not difficult to discern to which side of the argument the subject's own sentiments belong, but more reliable test cards as well as a more refined scoring scheme will have to be developed before the validity of the method can be definitely established.

An interesting variation of technique was developed by Jones and his co-workers, who, as part of their study on the social system in a factory town, investigated the attitudes of a cross sample of the population of Akron, Ohio, toward problems of "corporate property."⁵³ In interviews which lasted from thirty minutes to over an hour, Jones confronted his subjects with a series of carefully selected concrete stories which dealt with a clash of interest on economic and social topics and which were bound to produce a division of opinion. The stories, which generally evolved around the theme of a dominant company and its relation to the work force, always ended with the success of one of the antagonists, and the subjects were asked to state their degree of approval or disapproval of the action taken. The various responses of the subjects were weighted in terms of their revealed attitudes toward corporations. The results showed the existence of distinct attitude structures on the part of the various socioeconomic groups of the sample.

⁵² Murray and Morgan, *op. cit.*

⁵³ A. W. Jones, *Life, Liberty and Property* (Philadelphia: J. B. Lippincott Co., 1941).

Among the indirect attitude measurement techniques which have recently been developed, the "error-choice" method appears among the most promising. As originally conceived by Hammond, the technique attempts to relate the permanence of error in perception to the measurement of attitudes by provoking the subject to select pseudo-facts from memory and measuring the "direction" of the error.⁵⁴ Hammond did not claim to give a logically derived definition of attitudes, but emphasized the experimental nature of the problem by limiting his definition of attitudes to "those sources of energy, or affective states, capable of producing error in perception and recall." He applied the findings of the experimental literature on the permanence of the perceptive error and selective forgetting, and was able to develop two "attitude inventories" which significantly differentiated two criterion populations on questions of international and labor relations.

In an effort to verify the basic assumptions of the "error-choice" method, Weschler constructed an information test on labor-management relations, consisting of 40 questions, certain items of which were intended to elicit constant errors in the direction of the subject's known bias toward labor or management.⁵⁵ All of the items offered two possible answers; the 24 information items were definitely "real" in character and contained the correct choice as one of the alternatives, while the remaining 16 items were of the "error-choice" type, that is, were either controversial in character, or contained two incorrect answers as choices, and were of such a nature that the true answer was not easily accessible.

The test was administered to 186 advanced students at the University of California, Los Angeles. Two groups which were classified on the basis of their stated sympathy toward labor or management could be differentiated on the "pro-labor" attitude items of the scale. A statistical analysis of the data revealed 11 significant nonfactual "pro-labor" attitude items of varied discriminating ability. These items were weighted and utilized as the basis for rescoring the test. An analysis of the data, using the weighted scores, showed, among other results, that two groups with opposed stated sympathies toward labor or management differentiated themselves clearly on the "pro-labor" attitude portion of the inventory, that sex and age were not related to "pro-labor"

⁵⁴ K. Hammond, "Measuring Attitudes by Error-Choice: An Indirect Method," *Journal of Abnormal and Social Psychology*, Vol. 43 (January, 1948), pp. 38-49. This article is reproduced on pp. 170-82.—eds.

⁵⁵ I. R. Weschler, "An Investigation of Attitudes toward Labor and Management by Means of the Error-Choice Method, I," *Journal of Social Psychology*, Vol. 32 (August, 1950), pp. 51-62.

attitude, as measured by this inventory, and that, in a college population, the subject's income provided a significant clue to his attitudes toward labor or management.

In a follow-up investigation, which utilized a second form of the test inventory,⁵⁶ the method was further validated with the co-operation of a group of 67 management people, a group of 57 union leaders and members, and with 62 labor mediators who were included to serve as control. The results obtained showed the expected distribution of "pro-labor" attitude scores. Those individuals who held management jobs tended to score low on the "pro-labor" attitude items, while the union officers and members obtained significantly higher scores. The labor mediators who served as controls because they were assumed to be relatively neutral and whose scores were expected to distribute themselves around the mean of the total sample population generally scored in the "pro-labor" direction. It was found, however, that those mediators who were rated high by their colleagues in terms of their ability to do the job tended to score in the "neutral zone," that is, near the sample population mean; while those who were rated as "poor" by their colleagues scored either in the "pro-management" or in the "pro-labor" zones of the attitude range.

Recently two studies have been reported which indicate that the use of personal documents in the evaluation of attitudes is practically feasible and able to yield profitable results, as suggested by Allport.⁵⁷ The first of these studies involved a detailed analysis by a corps of psychologists and sociologists of over 200 manuscripts which were received in response to an announcement of a prize competition on the topic, "My Life in Germany Before and After January 30, 1933."⁵⁸ As a means of guiding and standardizing their work, these investigators employed a work schedule which dealt in detail with 18 different areas of interest, such as the development of attitudes toward National Socialism, sources of frustrations, identifications with groups, aggressive behavior, postmigration adjustment, and others. Wherever possible, use was made of rating scales and other quantitative devices and the results reported in statistical form.

⁵⁶ I. R. Weschler, "A Follow-up Study on the Measurement of Attitudes toward Labor and Management by Means of the Error-Choice Method: II," *Journal of Social Psychology*, Vol. 32 (August, 1950), pp. 63-69.

⁵⁷ G. W. Allport, *The Personal Document in Psychological Science*, Bulletin No. 4 (New York: Social Science Research Council, 1942).

⁵⁸ G. W. Allport, J. S. Bruner, and E. M. Jandorf, "Personality under Social Catastrophe," in C. Kluckhohn and H. A. Murray, *Personality in Nature, Society and Culture* (New York: Alfred Knopf, 1948), pp. 347-66.

The second illustrative study, as reported by Krech and Crutchfield,⁵⁹ was carried out by the Morale Division of the U.S. Strategic Bombing Survey in 1945. It was concerned with an analysis of captured German mail to ascertain various attitudes of people being subjected to strategic bombing. A random sample of mail originating in selected German cities was coded and rated for attitudes toward the Allies and toward the Nazi leadership, and for opinions about the course of the war, the probable outcome of the war, and so forth. The typical attitudes and opinions for a given community were then correlated with the objective facts about the extent and kind of bombing to which that community had been subjected by the Allies. By this analysis a number of important relationships between bombing and attitude were tentatively established. After cessation of hostilities, the findings were verified on the basis of direct and more intensive interviewing of a cross section of the German civilian population.

CONCLUSIONS

We have attempted to present a comprehensive review of the variety of "indirect" techniques which have been utilized for the measurement of attitudes and to describe some of the instruments which have shown themselves to be at least theoretically, if not yet practically, suitable for the task. In the majority of cases, it was impossible to determine the adequacy of these techniques because insufficient data have been reported, and the validity of the various methods remains to be tested by more and possibly better-controlled investigations. Nevertheless, there exists sufficient empirical evidence to indicate that this approach to the measurement of attitude expression is fruitful and valid.

The usefulness of any indirect method can be judged by the manner in which it approaches or fulfills the standards which have been empirically set up to define the "good" technique. Its range of applicability is the product of a set of circumstances, some of which are modifiable, and some of which are not. Thus, to a certain extent, the value of a given technique may depend upon the manner in which it is able to disguise its true purpose and can be adjusted to fit into a variety of different situations.

Another contributing factor to its usefulness may be its ability to produce a stimulus strong enough to force the subject to reveal attitudes which he has been interiorizing. Regardless of the type of materials used, their function must appear sensible to the respondent and

⁵⁹ Krech and Crutchfield, *op. cit.*, p. 239.

their administration controlled to avoid the influence and bias of the experimenter. The greatest need of all is the establishment of quantitative scoring and interpretive techniques which will make possible the introduction and rigid adherence to present-day scientific and statistical methods and allow for the rigorous investigation of reliability as well as validity.

Indirect methods of attitude measurement have established themselves as useful and provocative research tools, but their future status and development depend upon the type of research which will stress methodology rather than originality and scientific technique rather than ingenuity of design.

12. CAN THE CLINICAL TECHNIQUES BE VALIDATED?*

A psychologist himself, the author raises grave doubts concerning the validity of many of the psychological techniques, particularly projective tests, as applied to marketing research. Question is raised not only from a theoretical point of view, but on the basis of such empirical evidence as seems to be available as well. Of interest also are the author's comments regarding the research practices of research firms using psychological techniques.

As of this writing (Fall, 1957), the Motivation Committee of the Advertising Research Foundation is still in operation.

At a recent winter conference of the American Marketing Association, I attempted to estimate the promise and the limitations of clinical techniques as adapted to consumer motivation research. Here are some of the statements made at that time:

1. The long, informal type of interview . . . is not psychoanalysis.
2. [It is] not adapted to representative population surveys.
3. Our applications of projective techniques have been entirely empirical.
4. We must have more than testimonials to support projective techniques.
5. The conclusions of the clinical psychologist depend upon the clinician, himself, and the theories which he follows.
6. We can validate projective techniques only when . . . the measure can be applied to overt behavior.
7. Projective techniques have promise, but the most optimistic claim made for them is that they may predict with a correlation of $+ .30$.

* Adapted from a talk by Darrell B. Lucas, New York University, "Can the Clinical Techniques Be Validated?" given at the University of Illinois Marketing Symposium, 1955.

8. With regard to one "motivation study," [most of the report] deals with specific attitudes and behavior patterns.

9. Nearly all of the data [in this same report] could have been arrived at . . . by direct questions.

10. It has been my observation that . . . the people emphasizing "motivation research" . . . are often woefully lacking in fine field operations as well as in sampling methods.

These statements came from a talk covering limitations as well as the promise of clinically styled research techniques. Not all of the argument is so one-sided, but it must be admitted that there is great need for both procedural improvements and validation if those particular devices are to meet marketing research standards. This paper will discuss the problem of validating projective techniques. However, procedural questions cannot be allowed to cloud the issue of validity. Hence procedure will be discussed briefly, first.

PROCEDURAL ISSUES

With regard to the more definable procedures—if I may take the liberty of lumping the clinical methods together to save time—a few points stand out. Most applications are made in an atmosphere or a setting deliberately designed to give direction to the response. In a competitive type of survey, this atmosphere should not be allowed to give obvious aid to our client's property. Magazine studies should not be designed so as to minimize responses favorable to newspapers, television, radio, or other competing media. Sports magazine studies should not create a bias against home magazines or romance publications. And service magazine studies should not penalize those designed for entertainment or for widening our horizons.

With regard to the selection of field workers, their instruction, training, and supervision, it is essential to know what standards are being used. Perhaps the requirements for "clinical" methods are more rigid than for traditional surveys; on the other hand, perhaps field standards are of very little importance. One operator says you only have to pick workers who can get people to talk. Be that as it may—and it has no direct bearing upon the validity of the methods themselves—we should determine those standards and insist that they be maintained.

Most definable of the procedures in a survey is the sample size and design. Long, unstructured interviews are apparently unsuited to large population samples of probability design. Projective methods, at least, seem more capable of sampling success on a probability basis. So far there has been little evidence of representative sample selection

in publicized surveys of this type, and there has been some severe criticism aimed at crude sampling in motivation studies. In order to prove that rigid sample design is unnecessary, one operator offered to set up two small jobs, differing only in that one would use a probability sample and the other—shall we say—a more convenient selection. He would make all of the interviews and develop all of the conclusions for each study and compare them. The main objection to such a demonstration is that the responses to projective techniques are not tabulated, but are subjectively “interpreted.” How could he lose?

It is possible, of course, to develop certain tabulations based upon frequencies of responses, but these are usually obtained under conditions comparable to an open ballot—and often a heavily biased ballot. It is like having our elections based entirely upon an open ballot. I am sure that Ike would run far ahead of Eisenhower on any ballot requiring the memory and spelling of names. There must be enough objectivity in the reporting of projective survey tabulations to offset the lack of objectivity in the raw data.

Most other aspects of procedures using unstructured interviews and projective techniques are less easily definable. The methods used to develop the probing devices and the skill used in their design are not easy to specify. Obviously, the results cannot be valid if operators lack the skills for their proper use.

The Advertising Research Foundation has a committee now addressing itself to the validation of all methods used in motivation research. It has operated for three years under the continuous and able chairmanship of Dr. Wallace H. Wulfeck, chairman also of the executive committee of William Esty advertising agency. None of us can ignore the fact that clinical psychologists generally, and the Society for Projective Techniques in particular, are concerned with the same problem in its clinical setting. Their claims have been modest, indeed, and clinical experts have exhibited no shock or surprise by the harshness of things said on this subject of possible low validity. Psychologists, emphatically and generally, concede we are a long way from establishing even a medium level of validity for the projective approach.

Why, then, do we bother with procedures which hold so little scientific promise? Why does an unpaid advertising committee labor so long and hard to find the answer? Why do people continue to pay dues to a society named after such dubious techniques? The answer lies partly in the fact that clinical methods are not used exclusively for their scientific value or their guidance in prediction. In the process of diagnosis, most psychiatrists and staff psychologists rely upon what-

ever procedures help them in establishing clinical insights—insights which may defy explanation even by the doctors themselves. If a technique has the further advantage of offering a short cut, it is likely to gain all the more rapidly in favor.

This wandering off into technicalities is not intended to complicate our problem. The problem is complicated enough as it is. We are dealing with human motivation, and the discovery of the secrets of behavior is not as simple as the measurement of behavior itself. Any new tools which marketing people can devise or borrow for improving knowledge of motivation in any degree are worth examining. Any validity at all may well lead to measurable increases in marketing efficiency and will surely add to the competitive advantage of those who first can use such tools wisely.

One more preliminary consideration is the sharp contrast between clinical and advertising procedures. People get into clinics because of their behavior or misbehavior. By the time they get there, or as a first step in diagnosis, a finely detailed listing of behavior is prepared. The behavior comes first—then the clinic—then the application of such techniques as unstructured interviews and projective methods. In advertising, it is just the reverse. The advertiser aims to predict and control behavior which hasn't yet happened. In a sense, he has nothing to go on but the lessons learned from a more or less unrelated past. At best, he seems to have little chance of making as clear-cut discoveries from clinical methods as do the clinical users. Were it not that his decisions are so frequently forced upon him with little or no objective assurance, it is doubtful that the clinical methods would ever have been brought into the advertising scene.

VALIDATION STANDARDS

When we approach the question of validation in marketing and advertising, most of us probably think of validation in terms of overt behavior. The most obvious validating standard is buying behavior, and sales become our assumed measuring stick. But is this realistic? How much do we know about sales in relation to most national advertising? How many Cadillacs are sold by an advertisement headed "Class of 1985"? How many copy testers have succeeded in validating our traditional tests against a yardstick of sales? We cannot expect to validate the clinical type of research techniques by actual sales proof. We never proved the opinion method on the cash register, nor readership research, nor any of the other memory techniques. What, then, is to be the criterion?

Another gage probably guides the thinking of most of us with experience in established marketing research procedures. We judge our methods in terms of objectivity, meticulous control of field operations, precision of sampling, and rigid standards in the tabulating room and in report writing. Those have become our criteria for research into the unknown, confirmed sometimes by built-in checks against known bench-marks. Can the subjective interpretations arising from clinical types of studies be judged by these same standards? I don't know. When interpretation follows the pattern of some systematic psychology, the whole question of the validity of the system comes to the fore.

This probing of the behavior of perfectly normal consumers by unsophisticated field workers can hardly be compared with the work of skilled clinical psychologists exploring the troubled minds of highly motivated misfits. Nearly all of the requisite conditions of psychoanalysis are lacking in marketing applications of the unstructured interview. My own work in this field is in association with a practicing clinical psychologist who does only exploratory studies. Interviews are done by regular staff psychologists from mental clinics. Successive visits may total from three to thirty hours for the same consumer. There is no occasion for validation by quantitative methods since no projected estimates are offered.

Nevertheless, I am frequently asked to pass upon independently conducted "depth" studies which always seem to include projections and recommendations. Actually the projections are usually implied by the fact that recommendations are made for application to a national market. Frankly, these jobs don't give one much to go on. The samples are usually taken from a few concentrated spots. When a characteristic appears to be almost universal, it is, of course, convincing. Likewise, when it checks with one's own judgment, it carries added believability. But it isn't science, it isn't research, and it usually abides by rules made up to suit the occasion. We cannot admire the methods any more than we can validate them. We can, at times, admire the personal brilliance of the operator.

CLINICAL EVIDENCE

If we turn from consideration of unstructured interviews to the use of projective techniques, the problem appears a little less confusing. Projective methods usually follow a somewhat ordered procedure, and they center around printed forms and other physical aids which can be examined. Sampling does not meet with the obstacles so evident in prolonged, unstructured interviews. Projective methods often permit practical application of probability sample designs. There

is even some possibility of quantitative tabulation, although a subjective step is always involved in the interpretation. Altogether, there is a considerable basis for experimentation comparable with the more established survey procedures of marketing research. The possibility of statistical validation is not entirely ruled out.

A survey of the literature on evaluation of projective techniques in clinical and personnel psychology is not at all reassuring. The widely used Rorschach, after more than forty years, finds little experimental support no matter how well the procedure is standardized. One cannot, of course, standardize the personality of the administrator, the conditions and procedures under which the test is administered, or the processes of interpreting the results. Yet the Army Air Forces, employing the members of the Rorschach Institute to ensure competence and maximum uniformity, found that the Rorschach gave no promise as a predictive instrument for selecting trainee pilots. The correlation between tests during primary training and success as indicated by graduation was only $+0.04$.¹

Another application of the Rorschach was made by the Veterans Administration in connection with its program for training clinical psychologists. Using a success criterion of such factors as graduation, passing of licensing tests, and establishment of a clinical practice, the correlation was again but $+0.04$.² On the other hand, fellow trainees made predictive estimates at the $+0.25$ level, and the combined judgments of three staff members could attain a $+0.20$ level on the basis of a single day. Results of the Rorschach in medical student selection were similarly disappointing, even in separating the good ones from the poorest.³

Lest we appear guilty of condemning the Rorschach for its failures in applications to "sane" persons, instead of the maladjusted personalities for whom it was devised, we can cite evidence appropriate for that situation. Auld and Eron reported a study of the predictive value of the Rorschach in determining whether mental patients would continue undergoing treatment. They concluded that the Rorschach was not helpful, and that it was less indicative than the patients' own answers to direct questions.⁴ In the training of aviation pilots, where

¹ J. P. Guilford and J. I. Lacey, *Army Air Forces Aviation Psychology Program Research Reports Number 5*, 1947.

² E. L. Kelly and D. W. Fiske, *The Prediction of Performance in Clinical Psychology* (Ann Arbor: University of Michigan Press, 1951), p. 171.

³ L. D. Eron, "Use of the Rorschach Method in Medical Student Selection," *Journal of Medical Education*, Vol. 29 (May, 1954), pp. 35-39.

⁴ F. Auld, Jr., and L. D. Eron, "The Use of Rorschach Scores to Predict whether Patients Will Continue Psychotherapy," *Journal of Consulting Psychology*, Vol. 17 (April, 1953), pp. 104-09.

nervous disturbances are often the basis for elimination, Holtzman and Sells have reported similar failure of a group Rorschach combined with other projectives. The 50 highest rated cadets in a class of 1,504 were compared with 50 eliminated solely on the basis of nervous disturbances. The predictions were not significantly above a chance level, and one psychologist managed to steer a reverse course in excess of a chance level of error.⁵

Validation experiments involving the Thematic Apperception Test and sentence completion have resulted in almost as unimpressive correlations. In the experiment on prediction of performance in clinical psychology, the Thematic Apperception Test came out with $+0.11$ correlation, and the sentence completions with $+0.19$ correlation, compared with $+0.04$ for the Rorschach. None was at a useful level. In the instance involving 50 top aviation cadets versus 50 disqualified, nervous trainees, the use of sentence completions and other projectives failed to produce a useful predictive instrument in combination with the Rorschach.

It is not radical to conclude that these applications of projective methods have failed to demonstrate validity under experimental conditions. The failures have had to do with the testing of normal and abnormal personalities, especially in relation to specific job qualifications. That is not our marketing problem, but the evidence suggests caution in wholesale borrowing of these techniques for forecasting the behavior of American consumers.

Before leaving the atmosphere of the psychological clinic, in which the projective techniques were developed, it may be helpful to review other considerations. Why do the projectives have so much clinical use if they have so little validating proof? As one authoritative clinician said to me, "Why do they keep on using them? They don't correlate at all with psychoanalytical evidence and we all know it." But who has validated the predictive performances of psychoanalysis itself? Nobody has, and it is an almost universal fact that the psychoanalytical approach is not predictive. The criterion which my friend would use to validate the projectives has little more experimental scientific proof than do the projectives themselves.

I do not wish to seem critical of clinical psychology. My generalized training in that field has left a wholesome respect both for its accomplishments and for its progress. Even though the mental doctor

⁵ W. H. Holtzman and S. B. Sells, "Prediction of Flying Success by Clinical Analysis of Test Protocols," *Journal of Abnormal and Social Psychology*, Vol. 49 (October, 1954), p. 485.

seldom makes a prognosis, he does make a diagnosis which he may alter as he goes along. Projectives seldom become the final criterion for the classification or treatment of the personality. We have access to any number of testimonials to the value of projective methods. Those testimonials fail to satisfy experimental scientists, and they fail to satisfy the requirements of research. What, then, can be done about it?

Before attempting an answer, we must define the uses to which projective techniques are being put. It is clear that, in the clinic, projective methods are used to help find out what motivations are at work in the breakdown of an unadjusted personality. On the other hand, in the field of personnel selection, projectives have been tried for the purpose of discovering whether a normal personality pattern, and its underlying motives, give promise of future success. Both of these applications are quite different from marketing and advertising, where the problem is to see what can be discovered about the motives of consumers that will enable us to predict and influence their behavior in relation to specific products. We may also wish to test separate aspects of advertising and other marketing processes to determine their potential influence. We are dealing with normal people. We aren't looking for human faults and failures, but for motivational elements which can be identified and estimated in relation to their marketing effect. Moreover, in national advertising at least, we are anticipating sales and profit results which may not be appraised even after they have happened.

The implications of the situation seem clear. It is not likely that projective techniques can be validated against later effects in sales. We must either confine our validating tests to a limited choice of advertising and marketing situations or validate them against measurable, specific aspects of the whole process. Remember, too, that we are talking about the validation of a mixture of objective and subjective processes which can be only partially defined. We can proceed only within these limitations and develop evidence only to the extent that judgment permits us to accept it.

SOME EXAMPLES OF PROJECTIVE FINDINGS

Word association tests offer convenient examples of our problem of validation. We have been repeatedly told of the downfall of the word "prune." It has bad associations. But my waiter wasn't shocked this morning when I ordered prunes for breakfast. As another example, consider the word "buttery" from a recent advertisement by

the American Dairy Association. It came up with about the worst associations of any word in the copy of that ad. The words "smeary," "greasy," and "fatty" kept cropping up in our test. But what *are* the more pleasant synonyms for "buttery"? I guarantee you the dictionary won't help. On the other hand, don't we all have *nonverbalized* responses to "buttery" which are favorable? No other product succeeds in holding such a high competitive advantage on the basis of so little provable superiority. Yet I don't recommend the deliberate use of advertising words with bad associations. Association tests may help to discover bad associations, but it seems that only judgment can dictate what to do about it.

In projective tests involving pictures, a man may be shown in the cab of a powerful crane. Responses may include frequent references to the insecurity of the little man at the controls. Does this mean we should not use such a perspective in an industrial advertisement? Maybe not, but it seems to be a case where judgment should carry heavy responsibility. Perhaps this view of the machine is the means of giving the prospective buyer a maximum of information. Also, maybe most buyers are industrialists—not operators. Perhaps the industrialist is not particularly displeased to see his union operator looking a little insecure. Obviously, it requires consideration of many facts in order to arrive at a judgment in this matter. If projective techniques or unstructured discussions bring out such evidences of insecure feelings, we should use these tests. But often the information is of a preliminary nature rather than the basis of final decision.

VALIDATION ATTEMPTS

The Advertising Research Foundation, in co-operation with National Analysts, has spent considerable time developing a technique for studying motives and buying behavior for certain products. Drawings of shoppers were used for projective testing in the exploratory stages and produced data which were both stimulating and convincing. But gradually it was found that the projective approach could be abandoned for equally effective direct question forms, and at less cost. The counterargument has been offered that the projective devices provide a better atmosphere for the interview. That is a hard point to evaluate. Our own conclusion is that we will probably try projective methods in future explorations, but will replace them with more traditional question forms in the final job, if they are equally productive.

Currently, the Motivation Committee of the Foundation is sponsoring a validation study of the sentence completion technique for copy

testing. This committee has a keen interest in possible validation of all methods of the clinical type. Nevertheless, we have not yet designed a crucial experiment adequate for validating projective techniques against advertising performance or any relevant kind of overt human behavior. The sponsorship of the above test reflects the willingness to explore every avenue which may add more light. At least one other prominent committee and certain universities are seeking answers to the same problem.

In the meantime, the Motivation Committee is carrying on two other related activities. First, because motivation is only one aspect of consumer dynamics, the committee has taken over the buying behavior project and continued its development. Secondly, the committee has taken on the task of advising on new motivation studies and appraising published studies. There have been frequent meetings with operators and clients who are contemplating motivation studies, with the committee serving as a consultant. Committee thinking and approval have also gone into the one appraisal of a major media study already made available to Foundation subscribers. There seems little doubt that the standards of the forthcoming studies will be higher than the standards of some previous studies. It is hoped the validity will likewise increase.

There is one further point which may properly be reported on the views of Dr. Wulfeck's committee. There is a firm determination that studies involving indirect questions and projective techniques should avoid client bias and should observe rigid field standards the same as more orthodox field surveys. The argument that the good judgment and broad experience of the operator will make up for obvious shortcomings falls on deaf ears. It is in this area where most attempts at practical validation have so far fallen down. The methods could not be validated because the wholesale abandonment of operating standards obscured everything else; and, in one instance at least, the obvious bias of the interview design and looseness of final reporting were just as bad as the operations in the field.

There is much promise of better times ahead. Some research companies, even though specializing in the field of clinically styled techniques, are striving for standardization of interviews, rigid supervision of field work, and actual reporting of sampling success in relation to a probability design. Thus, the following paragraph was a prominent part of a research sales proposal involving projective methods:

Projective tests do offer us insights into personality structure that are often unobtainable by other regular survey means. However, there is also a danger because probing into the subconscious or unconscious, on the small scale that is

possible with a general survey of this type, gives us only a fragmentary view of a tremendous body of hidden or repressed motivations. When we see only a few pieces of the larger puzzle, we are in danger of making mistaken inferences about the parts we cannot see. Short of a complete psychoanalysis of respondents, we could not hope to obtain the kind of information we would need to construct an adequate picture of these deeply-buried motivations.

The encouraging signs on the horizon do not throw light on the validity of the methods, but they help make it possible to set up validating procedures. Heretofore, except for unpublished, private research, we have had little to validate even against the criterion of our own judgment. Operators employing methods adapted from the clinic have given us almost no information as to their specific methodology. Information revealed about the standards and process of application has usually indicated the presence of so many operating variables as to confuse the issue completely. The more studies that are conducted under definable conditions, the better the chances of validation. Even if the inherent, subjective elements in clinical methods continue to make clear-cut validation impossible, the well-ordered application of the methods will gradually lead to a more certain estimate of their true worth.

Some think science has already made its verdict, and that the verdict is negative. But I am not yet ready to go that far. It seems to me we should proceed on the following bases:

1. Continue to explore and to experiment with the clinical type of interview as one more source of new ideas, unique theories, and just plain hunches.

2. Put these tentative conclusions to the test by using established survey techniques for the purpose of measuring the extent and intensity of their occurrence in relation to an adequate sample of the intended market.

3. Continue to use the results of the clinical type of study as one more source of confirmation of conclusions arrived at by current copy tests and by other research methods not yet fully validated.

4. Try validating the clinical type of test against the rating methods we presently rely upon in the advertising business—readership ratings, recall scores, and so on. After all, if motives lead to the buying of products, motives also lead to the reading of advertisements and the clipping of coupon offers.

5. Finally, if the clinical methods continue to survive the current onslaught, work out a suitable experimental design for determining their validity.

13. PROJECTIVE TECHNIQUES FROM AN ANALYTICAL POINT OF VIEW*

Accepting the premise of the preceding article regarding the questionable validity of projective techniques as they have been applied, attention is now given to more explicit consideration of the methodological problems arising in the proper application of these techniques. From this point of view, many readers may be interested in the fivefold classification of why consumers behave as they do.

A psychiatrist, walking down the street one morning, happened to encounter an acquaintance walking in the opposite direction. The acquaintance nodded to him, and as they passed each other remarked, "Good morning." The psychiatrist stopped, looked after the passing figure, and muttered, "Now, what did he mean by that?"

This anecdote highlights, of course, the current tendency for many people to go overboard on the value and implications of psychology. Few today question the value of psychiatry, or of psychology, in explaining behavior patterns, but it by no means explains all actions or is necessarily the best of all alternatives in a particular instance. Much the same is true of projective techniques, or so-called "motivation research," as applied to economics and marketing. It has much to contribute to the understanding of consumer behavior, but the extent to which consumer behavior is explainable in terms of "basic" motivations is undoubtedly exaggerated.

ROLE OF PROJECTIVE TECHNIQUES

Projective techniques are generally regarded as involving the application and analysis of unstructured interviews to derive explanations of people's behavior. These interviews may involve long successions of probing questions extending over several hours (though not all depth interviews are projective in nature); they may rely on sentence completion tests, on Rorschach ink blot tests, on interpretation of pictorial representations, or on any number of other devices which researchers may concoct to induce the consumer to project his feelings.

Projective techniques are generally regarded as the vehicle by which so-called "motivation research" is accomplished, and many

* Adapted from a talk by Robert Ferber, University of Illinois, "Projective Techniques from an Analytical Point of View," given at the University of Illinois Marketing Symposium, 1955.

researchers currently seem to use the two terms synonymously both with each other and with respect to the determination of why people act as they do. In both of these respects, this is a grave mistake.

Projective techniques represent one method, or set of methods, of approach to the broad problem of determining "reasons why." It seeks to get at the answer to this problem by setting up a situation into which the consumer is induced to project his personal feelings and attitudes. This is not, however, the only means of getting at this problem and at times is not the correct method of doing so either. Direct questioning, mechanical observation, even such "prosaic" nonsurvey methods as time series analysis are at times powerful techniques for answering "why" questions. For example, why do farm families spend a smaller proportion of their incomes for medical care than do city families at corresponding levels of income (whatever that means)? One could approach this problem by interviewing samples of each group, by giving them sentence completions, ink blot tests, lie detector tests, or what have you. All sorts of differences may be uncovered, some psychological, some sociological, some anthropological, maybe even the correct answer—that the same services and facilities are not available to each group! (There is also the possibility that the sources of infection are different for each group.) This may or may not be uncovered by a survey, but even if it were, a much more sensible approach would have been to relate the distribution of medical facilities to consumer expenditures for medical care by area, which would have yielded the answer far more quickly and economically.

Perhaps a more pertinent example is the problem faced by the Survey Research Center some years ago of determining why some people bought savings bonds and others did not. Projective techniques in this instance might have determined, for example, that the members of one group tended to be more happily married than those of the other. But these techniques could hardly have determined a principal reason for the differences in purchases simply because the respondents themselves would not have realized it—that those who bought bonds were far more likely to have been approached personally than those who did not. It might well be that psychological factors were involved also, but as a basis for action—which is after all the *raison d'être* for "why" research—this was sufficient.

The manner in which projective techniques are related to these other approaches in determining the reasons why consumers behave as they do can be brought out by classifying these reasons with respect to consumer awareness of, and willingness to talk about, them. Doing

so, we find that the reasons for a particular action or attitude will fall into one or more of five possible categories, as follows:

First, the reason may be obtainable from the consumer, who knows the reason, and will tell it if asked. Here, direct questioning is all that would be needed if a survey were made on the subject. If consumers don't like the taste of a particular pudding, the reason why is readily ascertainable by direct questioning.

Second, the reason may be obtainable from the consumer, who knows the reason, but does not want to admit it. Such situations are likely to arise when the reason involves or implies social disapproval or ridicule, or may place the respondent in an unfavorable light. That many people may refuse to fly because they are afraid of heights; that some may refuse to consider buying a small car, such as the Nash Rambler, because of the sense of inferiority it gives them on the road; that some stout girls don't diet because they think they will never be asked out anyway, are not reasons likely to be derived from answers to direct questions. Where reasons of such a nature are suspected to exist, projective techniques are eminently suitable by allowing the respondent to express his feelings without embarrassing him and without direct questioning.

Third, the reason may be obtainable from the consumer, but he is not aware of it, because it relates to some underlying un verbalized social or psychological attitude. Projective techniques are ideally suited in such a problem, as illustrated by the example of why people did not buy instant coffee, namely, because of apparent identification of purchases of this "substitute" product with inferior housekeeping.

Fourth, the reason may be obtainable from the consumer; it relates to some behavioral, socioeconomic, or environmental characteristic, but he is not aware of it. The savings bond problem cited earlier is one such example. Another is one narrated by Alfred Politz to the effect that quick pickup in a car was found to be related to the softness of the accelerator pedal in the respondent's car, leading to the inference that the way to "increase" pickup, at least in the minds of consumers, was to soften the spring on the accelerator pedal rather than to increase horsepower. In such problems projective techniques are likely to be of little value because the consumer could not provide the necessary information consciously or otherwise.

Fifth, the reason is one which is not obtainable from the consumer by a survey technique. Thus, sales of a cereal may be low because the amount of display space allotted to it in stores may have been reduced. The income elasticity of housing expenditures by farm fami-

lies has been higher than that of urban families in the past largely because the farmer did not have as many varied outlets for spending money on other things as did urban families. In such problems, the study of aggregates or the compilation and analysis of nonconsumer data are likely to be most fruitful.

Thus, by this classification, projective techniques are of principal value in detecting two of five possible sets of "why" categories. This is no measure of their importance in relation to other techniques but is rather an indication of the type of situations to which they are most suitable. From a problem-solving point of view, the value of such a classification may be questioned, because a given problem is not likely to fall automatically into one of the above five categories; thus, the classification resembles to some extent a Bayes Theorem type of solution (given the probability attached to each event, the over-all probability is easily computed). But even such solutions can be of value. Basically, the selection of a proper technique is up to the insight and ingenuity of the researcher in formulating the problem, in recognizing alternative hypotheses, and in devising suitable test procedures. At the same time, however, *a priori* reasoning supplemented by suitable pretests and examination of preliminary data will often lead in a given problem to elimination of anywhere between one and four of the above five possibilities. At the least, it should serve to crystallize one's thoughts on the nature of the reasons for the observed actions and thereby aid in the selection of an optimal research technique.

METHODOLOGICAL PROBLEMS

The purpose of this section is to dwell on selected aspects of the methodological problem to which, on the basis of work in the area and review of the current literature, some attention is needed. No claim is made to comprehensive coverage either of breadth or of depth.

One aspect is the reliability of the interview itself. Some allege that a long, probing interview seeking to determine basic motivations is next to useless because it requires a trained analyst many months, if not years, of interviews with a person to ascertain his motivations. Since market research interviews are at best only a few hours in length and are generally not conducted by anyone with psychiatric training, how can such interviews be expected to get at inner motivations?

To many people, such an interpretation will seem unnecessarily harsh. But is it? In some ways, yes; in others, perhaps not. Basically

it would seem to depend on the problem at hand. The objective in most research studies is to ascertain the motivation for one particular action, or attitude, rather than the factors underlying one's over-all activities and outlook. For another thing, it is not necessarily the "basic" motivation that is sought in a research investigation but rather a motivation or factor that can serve as the basis for corrective action on the part of the client. Thus, if many people buy cigarettes on the basis of imagined connotation of brands with particular social groups, this is sufficient information for devising promotional literature. There is then no need to probe deeper for the motivations or forces underlying the reasons for these connotations, at least for purposes of understanding consumer behavior in this one respect.

In many cases, however, the basis for action may be a motivation so deep that only psychiatric methods could uncover it. This may well be the case, for example, in trying to determine why people prefer particular colors. In such a case, projective techniques obviously will not work, and the sooner this is found out (by pretests in most cases), the better it is for all concerned.

It is also perhaps self-evident that, if the objective is to get at the basic drives underlying consumer actions, much deeper analysis is required. At present, this area appears to be wide open for investigation, for relatively little work has been done as yet on an interdisciplinary project based on psychiatric interviews conducted from the viewpoint of explaining consumer buying motivations.

Question can also be raised regarding the necessity of having a trained analyst or clinical worker conduct the interviews. An interviewer with ordinary intelligence, a lot of common sense and, by no means least, a disarming and tactful personality can be fully adequate for the task with only a few hours' orientation on the objective of such interviews. Unfortunately, instances where interviewers receive as many as two or more hours orientation are probably the exception rather than the rule. One of the main flaws in most of the present commercial *modus operandi* is the inadequate orientation given to interviewers.

A related methodological problem is the interpretation of projective interviews. To what extent do research analysts overlook motivating factors or ascribe motivations to a person that are not justified by the interview report? The former is largely a matter of training, presumably, though impartial research on the matter is sadly lacking (perhaps for obvious reasons!). Ascribing motivations that do not exist is most easy to do in projective interviews. After scanning the

first few interview reports, most analysts will have developed certain hypotheses regarding motivating factors. With these preconceived notions in their minds, it can become quite easy to seek out substantiating evidence in the other reports and ignore other possible hypotheses.

Essentially, this danger of building a vested interest in a particular hypothesis exists in every research problem, but it is far more likely to be present in analyzing nonquantitative data where standards for evaluating responses are highly subjective. This is a problem which has received some attention in psychological journals but little attention from an economic, marketing, or statistical point of view, at least in the literature, although significant contributions to its solution would seem feasible by such approaches. These contributions would lie in formulating more meaningful approaches to the problem, rigorous formulations of hypotheses (particularly framing them in proper dimensions), procedures for validating interpretations of individual reports (such as a means of having the analyst check his interpretations with the interviewer supplemented by follow-up interviews where needed), and in various other ways.

The validity of the findings is perhaps the most widely discussed methodological problem in using projective interviews and has a number of facets in itself. One of these is the tendency to allege that whatever motivations may be found to be associated with particular behavior cause that type of behavior. It is an old pitfall in statistical analysis—though attracting new victims all the time—and is illustrated in statistical texts by such examples as the relationship between the stork population and birth rates, or the relationship between liquor sales and teachers' salaries. Since liquor sales and teachers' pay are closely correlated with each other in many areas, it obviously follows that the way to increase liquor sales is to boost teachers' pay—an experiment in which most academic people would be only too happy to participate.

Essentially the same pitfall is present in the analysis of projective interviews, and the fact that the relationships here are derived from individual cases rather than from aggregates does not remove this danger. Association between motivations and actions is just as likely to be mistaken for causation as is that between two statistical aggregates.

Furthermore, even if the relationship happens to be a causative one, the question of its pertinence arises. In other words, the relationship may be a very close one but the intensity of the relationship may be so low as to render it of little value for practical purposes. For ex-

ample, a close relationship may be uncovered between liking of crowds and attendance at concerts. Yet it may turn out that, so far as influencing or not influencing people to attend concerts is concerned, this factor is of little importance.

Question can also be raised of the extent to which the effect attributed to one variable in a two-variable relationship is actually due to some other factor either not included in the analytical framework or considered separately. For example, part of the relationship between reluctance to fly and fear of effect of an accident on one's family may be due to a relationship between the latter and one's financial position. Those who are in financial straits may be reluctant to fly or to take any imagined sort of physical risk lest an accident expose their position to the public as well as to the family. If this were true, the family-effect factor as a force in influencing tendency to travel by plane may be overemphasized.

Granted that these various dangers exist, what can be done about them? In this respect five suggestions may be advanced:

1. Approach the determination of a relationship as causative in much the same manner as porcupines are supposed to make love—very, very carefully! It is important to distinguish between different types of causes, particularly between causative factors that are sufficient in themselves and factors only operative in conjunction with other variables.

2. To avoid interviewer bias, which is far more prevalent in such studies than in fact-gathering surveys, use a number of different interviewers rather than rely on one or two. After the individual interviews have been interpreted, study those turned in by each interviewer as a group to ascertain the presence of distinctive patterns not present in the other groups.

3. Have at least two research analysts interpret at least part of the interviews independently of each other, and then compare results in the presence of a third analyst. The latter would be of general use as a buffer force in case of disagreements of opinion in addition to providing a broader perspective on the specific problem at hand.

4. Where causation does appear to exist, attempt to verify its existence by means of experimental design, utilizing a series of test patterns in selected areas.

5. If possible, test the validity of different techniques by employing them simultaneously on the same problem. Substantial progress in assessing validity would follow a series of tests applying two or three techniques—some projective, some of other types—to matched sub-

samples to solve the same problems, each technique under the supervision of one project director working independently of the others. Application of these techniques to several different problems—including, where competence permits, rotation of project directors among the techniques—would go a long way toward resolving the problem of assessing validity.

These suggestions by no means exhaust the possibilities, but perhaps they will serve to stimulate further thinking and empirical research in this area.

A final methodological point to be considered is the generality of the results of projective-interview studies. A common practice in many such commercial studies is to obtain somewhere between fifteen and fifty interviews, draw deductions from these interview reports, and then make recommendations for action on the explicit or implicit (usually the latter) assumption that the findings are valid for the entire population.

Except in certain rare cases, such an assumption would seem to be untenable. To see why, let us examine under what conditions this assumption *would* be tenable. Essentially, there are two broad sets of conditions for such a case. One is if the relevant variables or factors are almost perfectly correlated with each other. If all people who think of prunes as a symbol of intestinal discomfort do not buy prunes, and if all others do buy the product, there is clearly no sampling problem (assuming, of course, that some little bird tells you that such a magical variable exists in the first place). As few as two interviews would be sufficient for generalizing, and we are in the sampling statistician's Elysium of no worries of either representativeness or sampling error to disturb our reveries.

Clearly, however, such a condition is hardly likely to exist and even where it does, the birds one sees are hardly likely to exhibit or impart the supernatural analytical powers needed to make use of this fact. Hence, this brings us to the second set of conditions, which is that the sample members be representative of the population and that the sampling variation be small enough to yield operationally meaningful results. With the type of interviews that are conducted, however, the representativeness of the samples is obviously highly doubtful; and even if representativeness were secured, the sampling variation among fifteen to fifty interviews is likely to be considerable. Yet, if there is no representativeness and, quite likely, high sampling variation as well, on what basis can results of projective studies be generalized with greater assurance than results of any other sampling surveys?

The implicit basis on which this seems to be done is a feeling that motivational forces are not likely to change much from one person to another, so that once the motivations of a few people are known they can be attributed to all others as well. But such an assumption patently contradicts everything that is known about the psychological make-up of individuals. Psychologically as well as organically, an individual is a very complex animal. Differences between individuals are no less numerous with respect to psychological—and social psychological—characteristics than they are with respect to organic, economic, and social characteristics. To assume that the attitudes and motivations of a selected group represent those of the population is to ignore the existence of this variability and to present an unknowingly biased interpretation of the true state of affairs in the population.

This is true even for the simplest case, where dichotomies are involved. For if a relation is found between, say, purchase of prunes and one's attitude toward their laxative effects, the question arises of what proportion of purchasers and of nonpurchasers exhibit this attitude; without an adequate sample this cannot be determined.

The fact remains, therefore, that, in seeking reliable generalizations about a population, proper sampling methods must accompany the use of projective as well as of any other data-collection techniques. The main difficulties that will be immediately cited in the case of projective techniques are, of course, cost of obtaining a statistically adequate number of interviews and the matter of securing co-operation from those approached. The latter is not much of a problem if a correct approach is used—a common failing of many surveys in this respect is to use the same approach to all sample members irrespective of neighborhood, level of income, and so on—and if high-caliber interviewers are employed.¹

The cost factor, however, is indeed a problem and one that has to be faced squarely by the survey organization. Frank discussion of this factor with the client, in the case of commercial research, with indications of the consequences that can result from an inadequate sample, will often lead to the appropriation of the additional money. At times, however, it will not, and in such situations it seems best not to proceed rather than to have inadequate funds and run a large risk of coming up with unreliable results.

¹ The allegation made at times that representativeness cannot be secured because of the difficulty of securing co-operation from all segments of a population is not necessarily borne out by experience. See F. H. Sanford and I. M. Rosenstock, "Projective Techniques on the Doorstep," *Journal of Abnormal and Social Psychology*, Vol. 47 (January, 1952), pp. 3-16.

One other possibility that can at times reduce cost and still yield reliable results is to use projective techniques on a small sample, analyze the results, and derive hypotheses for testing by means of a larger survey but with less intensive interviews.

SUGGESTIONS FOR FUTURE USE

A few suggestions are offered with regard to the use and further application of projective techniques.

1. Projective techniques constitute a valuable addition to the tools available for the study of consumer behavior. Enthusiasm over this tool, however, has led to its misuse in many instances and to lack of recognition of the fact that it is only one of many tools for determining why consumers act as they do.

2. Projective techniques will provide the correct answers to only a limited number of "reasons why" problems and by no means to all of them. There are situations where either projective techniques or some other approach will yield essentially the same results. In view of the higher cost of the former in most such instances, it is doubtful whether as much preference should be given to the projective methods as has been the case in the past.

3. Where projective techniques are not employed as the principal approach in a particular problem, they nevertheless can frequently prove highly useful for pretest purposes. A few interviews of this type will often provide valuable material for formulating hypotheses with greater precision and for phrasing questions more meaningfully. The results of projective interviews often represent a beginning to a problem rather than an end in itself.

4. Adherence to established sampling principles is of as much importance in the application of projective techniques as it is in the use of any other data-collection technique.

5. The familiar socioeconomic classifications invariably serve as the basis for stratifying samples interviewed by projective techniques. In many instances, however, psychological bases of stratification would seem to be far more pertinent, yet they are generally ignored.² Such means of stratification do, of course, raise some difficult problems—particularly of sample representation and of estimation of population sizes—but conceptually they are no more difficult than those faced originally in setting up the currently used bases of stratification. The

² On the other hand, this problem has not been entirely overlooked. See, for example, C. McGuire, "Social Stratification and Mobility Patterns," *American Sociological Review*, Vol. 15 (April, 1950), pp. 195-204; and P. K. Hatt, "Stratification in the Mass Society," *American Sociological Review*, Vol. 15 (April, 1950), pp. 216-22.

sooner one makes a beginning at these problems, the sooner they will be solved.

6. It would be of great value to progress in this area, as it is for other subjects, if instances of unfavorable results from the use of projective techniques were reported as well as the favorable ones. This would perhaps be too much to expect—for obvious reasons—but like Hollywood's use of superlatives, the flooding of the literature with so many success stories tends after a while to produce an effect opposite to that intended. Just about everyone knows that quite a few of these studies fail to produce meaningful results—in some cases, because of faulty application; in others, through perhaps no fault of the researchers involved. Pooling and discussion of unsuccessful studies can produce a far better understanding of the drawbacks and limitations of these techniques than far more numerous accounts of successful, and often exaggerated, exploits.

7. A more critical approach is needed to the question of validity of projective interview study results. This applies to the interpretation of individual interview reports, to the distinction between causation and association, and to the verification of the results. With regard to all of these aspects, the complementary use of controlled designs, even simple ones, could be of considerable value. With only a little additional effort, the use of such a combination of techniques would lead to the rapid accumulation of a substantial body of knowledge on validation procedures.

If one thing is evident from this discussion, it is that the state of projective techniques is such that an objective and experimental approach to the subject can be highly fruitful and is also badly needed. It is only through such an approach that the usefulness, as well as the limitations, of these techniques can be established and gain wide acceptance as a tool for studying human behavior.

14. THE THEMATIC APPERCEPTION TEST*

The projective technique that has received most attention in market motivation research is the Thematic Apperception Test (TAT), in which

* By Gardner Lindzey, University of Minnesota. The original title and source of publication of this article is "Thematic Apperception Test: Interpretive Assumptions and Related Empirical Evidence," *Psychological Bulletin*, Vol. 49 (January, 1952), pp. 1-25.

the respondents are asked to construct stories about a standard set of pictures. Though consulting psychologists claim that TAT is not used in the same sense in market research as in clinical work, anyone who wants to use TAT or bases decisions on results obtained from TAT is well advised to become acquainted with the concepts underlying this test and problems involved in interpreting the data.

The present article contains much invaluable information on this subject and brings out clearly the many difficulties encountered in interpreting TAT data and the many questions surrounding the validation of this test.

The chief purpose of this article is to state the assumptions customarily involved in interpreting the Thematic Apperception Test (TAT),¹ and to examine the logical considerations and some of the empirical evidence that can be used to verify or reject each of these assumptions.

Aside from certain historical ties to psychoanalysis, the theoretical and empirical continuity between projective testing and the remainder of psychology has been a subject of little interest. This is true in spite of an increasingly large amount of research and formulation in other areas of psychology that is directly pertinent to the activities involved in projective testing. In particular, the research of the "new-look" perceptionists represents an important and fertile link between projective testing and more traditional domains of psychology. The initial outline of such a continuity has been traced by Blake and Wilson in a study of the influence of depressive tendencies upon selectivity in Rorschach response,² Bruner in a discussion of the Rorschach test,³ Douglas in an investigation of temporal factors in perception,⁴ Siipola in an examination of the effect of color upon Rorschach response,⁵ and Stein in an ingenious study employing tachistoscopic exposure of Rorschach stimulus material.⁶ All of these studies illustrate the feasibility and fruitfulness of relating perception findings to material elicited by Rorschach-like techniques. One of the aims of the present paper is to stress the desirability of relating such theory and research

¹ These same assumptions are customarily employed in connection with other story-construction projective techniques, such as Make-A-Picture-Story Test, Four-Picture Test, and Tri-Dimensional Apperception Test.

² R. R. Blake and G. P. Wilson, Jr., "Perceptual Selectivity in Rorschach Determinants as a Function of Depressive Tendencies," *Journal of Abnormal and Social Psychology*, Vol. 45 (July, 1950), pp. 459-72.

³ J. S. Bruner, "Perceptual Theory and the Rorschach Test," *Journal of Personality*, Vol. 17 (December, 1948), pp. 157-68.

⁴ A. G. Douglas, "A Tachistoscopic Study of the Order of Emergence in the Process of Perception," *Psychological Monographs*, Vol. 61 (1947), No. 6, Whole No. 287.

⁵ E. M. Siipola, "The Influence of Color on Reactions to Ink Blots," *Journal of Personality*, Vol. 18 (March, 1950), pp. 358-82.

⁶ M. I. Stein, "Personality Factors Involved in the Temporal Development of Rorschach Responses," *Rorschach Research Exchange*, Vol. 13 (December, 1949), pp. 355-414.

to the findings of investigators who employ the Thematic Apperception Test.

One may legitimately object to appraising empirically statements treated as given, or axiomatic. It is not clear, however, that all psychologists would concur in giving axiomatic status to these statements. As is true of so much of psychological formulation, the distinction here between analytic and empirical is not clearly delineated—what is one person's empirical generalization is another's axiom. Further, most psychologists probably agree with MacCorquodale and Meehl in subscribing to a methodological position that emphasizes the use of *only* analytic constructs that interact smoothly with available empirical knowledge.⁷ Consequently, examination of these assumptions in the light of empirical evidence seems justified on the one hand by the questionable axiomatic status of the assumptions and on the other by the general acceptance of psychologists that axioms should not violate observational data.

The assumptions to follow vary greatly in their generality. In fact, they are logically related only by virtue of their frequent use in the interpretation of material secured from one particular type of projective technique. They do not represent all of the assumptions employed in making such interpretations. Included are only those not sufficiently general to be common among all psychologists and yet, within the prescribed situation, sufficiently common so that almost any individual engaged in this activity would employ them. For the latter reason, attention is not given here to the assumptions involved in "formal" or "sign" analysis of story projective material where the meaning, or the thematic qualities, of the material is minimized, for example, Balken and Masserman⁸ and Wyatt.⁹ Nor has any attempt been made to explore the assumptions involved in the more recent "sequence analysis" of projective responses.¹⁰

THE ASSUMPTIONS AND RELATED EVIDENCE

The assumptions are divided into three crude groups. First is the most general assumption, fundamental to all projective testing. Second

⁷ K. MacCorquodale and P. E. Meehl, "On a Distinction between Hypothetical Constructs and Intervening Variables," *Psychological Review*, Vol. 55 (March, 1948), pp. 95-107.

⁸ E. R. Balken and J. H. Masserman, "The Language of Phantasy: III. The Language of the Phantasies of Patients with Conversion Hysteria, Anxiety State, and Obsessive-Compulsive Neuroses," *Journal of Psychology*, Vol. 10 (1940), pp. 75-86.

⁹ F. Wyatt, "Formal Aspects of the Thematic Apperception Test," *Psychological Bulletin*, Vol. 39 (July, 1942), p. 491 (abstract).

¹⁰ M. B. Arnold, "A Demonstration Analysis of the TAT in a Clinical Setting," *Journal of Abnormal and Social Psychology*, Vol. 44 (January, 1949), pp. 97-111.

are those assumptions concerned with procedures employed in determining the diagnostically significant portions of the fantasy productions. Third are those assumptions involved in relating the significant portions of the protocols to other forms of behavior.

Primary Assumption:

In completing or structuring an incomplete or unstructured situation, the individual may reveal his own strivings, dispositions, and conflicts.¹¹

Assumptions Involved in Determining Revealing Portions of Stories:

1. In the process of creating a story the story-teller ordinarily identifies with one person¹² in the drama, and the wishes, strivings, and conflicts of this imaginary person may reflect those of the storyteller.

a. It is assumed further that the identification figure can be established through the application of a number of specific criteria, such as person appearing first in the story, person doing most of the behaving, and person most similar to storyteller.

b. It is also assumed that additional figures in the stories such as father, mother, or brother often may be equated to the real-life counterparts of the storyteller and the behavior of the hero toward them used as indicative of the storyteller's reactions to these persons.

2. The storyteller's dispositions, strivings, and conflicts are sometimes represented indirectly or symbolically.

3. All of the stories that the subject creates are not of equal importance as diagnostic of his impulses and conflicts. Certain crucial stories may provide a very large amount of valid diagnostic material while others may supply little or none.

4. Themes or story elements that appear to have arisen directly out of the stimulus material are less apt to be significant than those that do not appear to have been directly determined by the stimulus material.

5. Themes that are recurrent in a series of stories are particularly apt to mirror the impulses and conflicts of the storyteller.

Assumptions Involved in Deriving from Revealing Portions of Fantasy Material Inferences about Other Aspects of Behavior:

1. The stories may reflect not only the enduring dispositions and conflicts of the subject, but also conflicts and impulses that are momentarily aroused by some force in the immediate present.

a. The further assumption is frequently made that both the enduring and temporary processes are reflected in stories in the same manner.

2. The stories may reflect events from the past that the subject has not himself actively experienced, but rather has witnessed or observed, for example, a street scene, story, or motion picture.

¹¹ The terms "strivings, dispositions and conflicts" are meant to designate all the attributes or aspects of the person that the clinician is interested in or wishes to measure. One could readily add to this list such terms as: personality organization, primitive fixations, complexes, needs and press, or any others that seemed necessary to represent those aspects of the person that are being explored.

¹² Most investigators accept the possibility of multiple identifications. This necessitates the same assumption, however, and only complicates somewhat the general problems of establishing interpretive rules.

a. It is assumed further that, although the subject has not himself experienced these events, and is telling them as he observed them, the fact that he selects these events, rather than others, is in itself indicative of his own impulses and conflicts.

3. The stories may reflect group-membership or socio-cultural determinants in addition to individual or personal determinants.

4. The dispositions and conflicts that may be inferred from the storyteller's creations are not always reflected directly in overt behavior or consciousness.

It should be emphasized that the research to be discussed in connection with the various assumptions can not serve as a validation of this kind of projective testing. At most, it can demonstrate that the very general assumptions lying behind this kind of activity are not in direct conflict with available empirical findings. This may increase the "plausibility" of projective interpretations, but the task of demonstrating the utility of specific rules of interpretation remains.

It is likewise important to realize that this article omits most of the pertinent empirical evidence resulting from clinical use of this technique. The omission of this idiographic research or observation is not the consequence of any feeling that such data are not of immense value. Rather, in an article emphasizing the relation of projective testing to the remainder of psychology, it seemed desirable to stress most heavily those kinds of research that come closest to meeting the experimenter's demand for empirical control and intersubjectivity of method. Thus, most of the evidence to be referred to has been secured with some attempt at maintaining adequate empirical controls.

DOES THE INDIVIDUAL REVEAL HIS OWN DISPOSITIONS AND CONFLICTS IN COMPLETING AN UNSTRUCTURED SITUATION?

The assumption that an individual reveals his own dispositions and conflicts in completing an unstructured situation is not limited to story-construction tests but lies at the heart of all projective testing. Fortunately, in view of its ubiquity, there is a host of very general experimental verification, some drawn from laboratory research.

One of the first investigators to point to the relationship between motivational factors and response in an unstructured situation was Sanford. He demonstrated in a series of experiments that the food responses of subjects varied as a function of the amount of food deprivation they had undergone. The responses were given in the act of telling stories to ambiguous pictures, making word associations, and in other situations where the stimulus was sufficiently unstructured to

permit either food or nonfood responses.¹³ This same general function was later demonstrated under somewhat different conditions by Levine, Chein, and Murphy,¹⁴ and McClelland and Atkinson.¹⁵ The results of these three investigations do not agree in detail, but all were able to demonstrate some kind of variation in response to ambiguous stimulus situations as a function of food deprivation.

In an early study by Murray, it was shown that estimates of "maliciousness" of faces in photographs varied directly with an experimentally induced state of fear. After a fear-producing game, children rated a series of photographed faces as significantly more malicious than they had rated the faces previous to the game.¹⁶ A recent study by Katz has shown that the manner in which an individual completes an incompletely drawn face varies systematically with the kind of experience he has undergone previous to the activity. Thus, a group that had just failed on a test of reasoning ability differed significantly from a group that had just spent their time in a neutral activity. Presumably the manner in which the individual completed this unstructured situation was directly related to the motivational or emotional conditions aroused by the experimental treatment.¹⁷

A series of investigations by Bruner and Postman have demonstrated that, when stimulus material is made ambiguous by brief tachistoscopic exposure, the responses or "guesses" of the subjects vary systematically with motivational variables.¹⁸ These same investigators, as a result of their perception research, have come to occupy a much more extreme position than that implied by the above assumption. They suggest that:

Most experimenters who have worked with need and attitude factors in perception have assumed, sometimes quite explicitly, that only in highly equivocal

¹³ R. N. Sanford, "The Effects of Abstinence from Food upon Imaginal Processes: A Preliminary Experiment," *Journal of Psychology*, Vol. 2 (1936), pp. 129-36; "The Effects of Abstinence from Food upon Imaginal Processes: A Further Experiment," *Journal of Psychology*, Vol. 3 (1937), pp. 145-59.

¹⁴ R. Levine, I. Chein, and G. Murphy, "The Relation of the Intensity of a Need to the Amount of Perceptual Distortion: A Preliminary Report," *Journal of Psychology*, Vol. 13 (April, 1942), pp. 283-93.

¹⁵ D. C. McClelland and J. W. Atkinson, "The Projective Expression of Needs: I. The Effect of Different Intensities of the Hunger Drive on Perception," *Journal of Psychology*, Vol. 25 (April, 1948), pp. 205-22.

¹⁶ H. A. Murray, "The Effect of Fear upon Estimates of the Maliciousness of Other Personalities," *Journal of Social Psychology*, Vol. 4 (August, 1933), pp. 310-29.

¹⁷ I. Katz, "Emotional Expression in Failure: A New Hypothesis," *Journal of Abnormal and Social Psychology*, Vol. 45 (April, 1950), pp. 329-49.

¹⁸ J. S. Bruner and L. Postman, "An Approach to Social Perception," in W. Dennis (ed.), *Current Trends in Social Psychology* (Pittsburgh: University of Pittsburgh Press, 1948), pp. 71-118.

stimulus situations can such "nonsensory" factors operate. . . . But all stimulus situations are potentially equivocal and cease to be so only to the extent that selection, accentuation, and fixation have taken place. Perception occurring without the contribution of such adaptive factors is as unthinkable as perception without the mediation of receptive nerve tissue. . . . *Adaptive factors in perception are not limited to unstable stimulus situations.*¹⁹

In light of this evidence, the assumption that motivational factors are revealed in completing unstructured situations seems clearly warranted.

DOES THE SUBJECT IDENTIFY WITH SOME FIGURE IN THE STORIES HE TELLS AND CAN THIS FIGURE BE ESTABLISHED RELIABLY?

There seem to be three modal positions in regard to identification in the story-telling process. First, we may make the assumption, indicated above, that there is *ordinarily* a single identification figure in each story. This assumption can be complicated greatly by a number of special conditions of the type already suggested by Murray.²⁰ Second, we may assume a continuum of identification with those figures in the stories that are very similar to the storyteller possessing a maximum of the subject's attributes and those that are very dissimilar possessing a minimum. This is the assumption made by Sears, who, working primarily with doll techniques and heavily influenced by the notion of "stimulus generalization," has suggested specifiable dimensions along which the degree of identification can be expected to vary. Thus, in doll play, figures resembling the subject in age and sex will show more characteristics of the subject than those figures that are dissimilar in age and sex.²¹

Third, we might, with Henry,²² Piotrowski,²³ and others, make no attempt to locate the hero but simply look on all characters in the constructed stories as representative of aspects of the storyteller. This last alternative is perhaps the least happy, not only because a large amount of diverse clinical experience militates against it, but also because it leads to certain drastic limitations upon the diagnostic use of

¹⁹ J. S. Bruner and L. Postman, "Tension and Tension Release as Organizing Factors in Perception," *Journal of Personality*, Vol. 15 (June, 1947), pp. 301, 307.

²⁰ H. A. Murray, *Thematic Apperception Test Manual* (Cambridge, Massachusetts: Harvard University Press, 1943).

²¹ R. R. Sears, "Effects of Frustration and Anxiety on Fantasy Aggression," *American Journal of Orthopsychiatry*, Vol. 21 (July, 1951), pp. 498-505.

²² W. E. Henry, "The Thematic Apperception Technique in the Study of Culture-Personality Relations," *Genetic Psychology Monographs*, Vol. 35 (February, 1947), pp. 3-135.

²³ Z. A. Piotrowski, "A New Evaluation of the Thematic Apperception Test," *Psychoanalytic Review*, Vol. 37 (April, 1950), pp. 101-27.

the TAT. If we adopt this assumption, we are more or less forced to give up the attempt to appraise the subject's attitudes toward other persons. Piotrowski suggests that we eliminate the hero and assume "that every figure in the TAT stories expresses some aspect of the testee's personality,"²⁴ while somewhat later he suggests that many stories "reflect what the subject thinks and feels about persons represented by the TAT figures, i.e., about the old and the young, the male and the female."²⁵ Thus, with no rules for differentiating, we are told that all figures represent characteristics of the hero, but also that the characteristics of some figures represent the story-teller's attitudes toward other persons. If we assume that all figures are equally representative of the story-teller's characteristics, either we must give up attempts to appraise the subject's attitudes toward other persons through this instrument or else engage in some kind of dialectic in the effort to defend such attempts.

There is, as yet, no unequivocal answer to the question of which of the above assumptions is most useful. In order to show that the assumption of a single identification figure is warranted, it is necessary to demonstrate relationships between the behavior of the subject and the imaginary behavior of the hero that do not exist between the behavior of the subject and the behavior of nonhero figures. If this difference is not shown, one can always suggest that the sensitivity of the test to aspects of the subject is a result, not of any specific identification process, but rather of a generalized reflection of motivational state of the sort that Murray,²⁶ Bruner and Postman,²⁷ and others have reported. Thus, even if *all* figures in the stories represented aspects of the storyteller, we might arbitrarily decide to use only a certain percentage of these figures as diagnostic of the story-teller's attributes, that is, employ the hero and nonhero distinction, and we would still expect the test to show some sensitivity to variations in storyteller behavior even though we were wasting much of its power. Just as the subject's motivational states are reflected in free association, or in the presolution hypotheses of tachistoscopic response, so also the verbal flow accompanying the storytelling process may mirror motivational states without any intervening identification with particular actors in the stories.

²⁴ *Ibid.*, p. 107.

²⁵ *Ibid.*, p. 113.

²⁶ Murray, "The Effect of Fear upon Estimates of the Maliciousness of Other Personalities," *op. cit.*

²⁷ Bruner and Postman, "An Approach to Social Perception," *op. cit.*

The distinction between "heroes" or identification figures on the one hand and "nonheroes" on the other is greatly complicated by the fact that the storyteller's orientation or attitudes toward other persons is intimately related to his own psychic make-up. Thus, we might mistakenly employ an identification figure as representative of the storyteller's attitudes toward other people and actually secure considerable information about the individual's external orientation because of the similarity between "self" and "perceived other." Or, conversely, we might take a figure intended by the storyteller as an "other" object and find mirrored in it much of real pertinence to the storyteller's own personality structure.

What is needed to answer the above question is a clear demonstration of more intimate relations between "hero" behavior or attributes and the storyteller than can be shown between "nonhero" behavior and the storyteller. Results, such as Bellak's, showing that the frequency of aggressive words rises when the individual is frustrated, do not demonstrate that the identification figure is necessarily displaying more aggressive behavior.²⁸ Nor are studies satisfactory that simply demonstrate variations in "hero" activities or attributes that relate to behavior or attributes of the storyteller. Even if Bellak had shown that aggressive behavior on the part of heroes rose following frustration, this would still not provide the needed information. In this case it is quite possible that analysis of the behavior of the nonidentification figures would reveal the same increase in aggressive behavior, indicating not identification, but simply an increase in verbal responses pertinent to the need or conflict in question.

Lindzey has demonstrated that, following a social frustration, the incidence of aggressive acts in TAT stories carried out by heroes against others increased more than the incidence of aggressive acts carried out by "other" figures.²⁹ If the incidence of the two kinds of acts is combined, the resulting shift is greater than the shift in either "self" or "other" figures alone. These findings are complicated by the fact that the frustration situation was of such a nature that both extrapunitive responses and a view of the environment as hostile and threatening would be expected to rise following the experimental treatment. Thus, the increase in aggressive responses on the part of non-hero figures might be considered a result of the increase in aggressive

²⁸ L. Bellak, "The Concept of Projection: An Experimental Investigation and Study of the Concept," *Psychiatry*, Vol. 7 (November, 1944), pp. 353-70.

²⁹ G. Lindzey, "An Experimental Examination of the Scapegoat Theory of Prejudice," *Journal of Abnormal and Social Psychology*, Vol. 45 (April, 1950), pp. 296-309.

tendencies on the part of this storyteller or a reflection of the fact that the storyteller viewed the external world as a more hostile and threatening place. It is clear that what is needed is an experimental treatment where, given the subject's identification with a hero-figure, the predictions to be made for "hero" and "other" figures will either be opposed or widely different.

Accepting the desirability of locating identification figures, can these figures be specified precisely? Several studies in which relatively high interscorer reliability coefficients were secured imply that it is possible to obtain reliable agreement among different scorers in establishing the identification figure.³⁰ Presumably, a positive correlation for need and press ratings could be secured only if the different raters were treating the same figures as hero. In addition, Mayman and Kutner report complete agreement in 89 per cent of the cases between two raters who independently determined the identification figure for a series of 91 stories.³¹

In general, then, the feasibility of the identification assumption cannot be clearly demonstrated at present, although empirical evidence suggests that identification figures can be established with reasonable reliability.

ARE IMPULSES AND CONFLICTS OF THE SUBJECT SOMETIMES REPRESENTED SYMBOLICALLY?

The clinical use of the concept of symbolism is intimately related to a number of psychological concepts employed by more rigorous and empirically oriented investigators. All behavior theorists have encountered the thorny problems posed by interchangeability or substitutability in behavior of stimuli and responses. These problems have led to the formulation of a number of psychological concepts including displacement, substitution, stimulus generalization, and vicarious mediation.

Apparently most psychologists agree that when a response is interfered with either by internal or external barriers it will frequently become altered to another form or else be directed toward a new object. There is a large body of research, much of it carried out on

³⁰ R. Harrison and J. B. Rotter, "A Note on the Reliability of the Thematic Apperception Test," *Journal of Abnormal and Social Psychology*, Vol. 40 (January, 1945), pp. 97-99; S. S. Tomkins, *The Thematic Apperception Test* (New York: Grune and Stratton, 1947).

³¹ M. Mayman and B. Kutner, "Reliability in Analyzing Thematic Apperception Test Stories," *Journal of Abnormal and Social Psychology*, Vol. 42 (July, 1947), pp. 365-68.

animal subjects, demonstrating this relationship and exploring some of the conditions under which it operates.³² Of the animal investigations, perhaps the most pertinent to our discussion is Miller's demonstration that a learned aggressive response can be generalized or displaced from an original stimulus object (another rat) to a substitute stimulus (doll) when the first stimulus is made unavailable.³³ This process can be considered a rough paradigm of symbolic representation.

A special question of considerable importance to the projective tester is whether symbolic transformations can, and customarily do, take place without awareness of the subject. Most of the clinical studies discussed below deal with symbolic representations of which the subject was not aware initially and of which he could become aware only under special conditions, such as interpretation or therapeutic change.

This capacity of the organism to engage in symbolic representation without awareness of the process has been demonstrated under better-controlled circumstances by other investigators. Diven and Haggard have both shown that a word may become a substitute stimulus for an electric shock and evoke a galvanic skin response that differentiates it from "non-shock" words, even where the subject is unaware of the connection between the word and the electric shock.³⁴ In similar fash-

³² J. S. Brown, "The Generalization of Approach Responses as a Function of Stimulus Intensity and Strength of Motivation," *Journal of Comparative Psychology*, Vol. 33 (April, 1942), pp. 209-26; C. N. Cofer and J. P. Foley, Jr., "Mediated Generalization and the Interpretation of Verbal Behavior: I. Prolegomena," *Psychological Review*, Vol. 49 (November, 1942), pp. 513-40; L. Grandine and H. F. Harlow, "Generalization of the Characteristics of a Single Learned Stimulus by Monkeys," *Journal of Comparative and Physiological Psychology*, Vol. 41 (October, 1948), pp. 327-38; C. L. Hull, *Principles of Behavior: An Introduction to Behavior Theory* (New York: Appleton-Century, 1943); C. L. Hull, "The Problem of Primary Stimulus Generalization," *Psychological Review*, Vol. 54 (May, 1947), pp. 120-34; K. S. Lashley and M. Wade, "The Pavlovian Theory of Generalization," *Psychological Review*, Vol. 53 (March, 1946), pp. 72-87; G. Razran, "Stimulus Generalization of Conditioned Responses," *Psychological Bulletin*, Vol. 46 (September, 1949), pp. 337-65; G. Razran, "Attitudinal Determinants of Conditioning and of Generalization of Conditioning," *Journal of Experimental Psychology*, Vol. 39 (December, 1949), pp. 820-29; D. D. Wickens, "Stimulus Identity as Related to Response Specificity and Response Generalization," *Journal of Experimental Psychology*, Vol. 38 (August, 1948), pp. 389-94.

³³ N. E. Miller, "Theory and Experiment Relating Psychoanalytic Displacement to Stimulus-Response Generalization," *Journal of Abnormal and Social Psychology*, Vol. 43 (April, 1948), pp. 155-78.

³⁴ K. Diven, "Certain Determinants in the Conditioning of Anxiety Reactions," *Journal of Psychology*, Vol. 3 (1937), pp. 291-308; E. A. Haggard, "Experimental Studies in Affective Processes: I. Some Effects of Cognitive Structure and Active Participation on Certain Autonomic Reactions during and following Experimentally Induced Stress," *Journal of Experimental Psychology*, Vol. 33 (October, 1943), pp. 257-84.

ion, McCleary and Lazarus showed that words to which a galvanic skin response had been conditioned through the use of electric shock could elicit the galvanic skin response even when exposed tachistoscopically at such rapid speeds as to prevent conscious recognition.³⁵

Certainly, since Freud, the clinical interpretation of diagnostic material, whether projective or not, has depended heavily upon the assumption of symbolic transformations. There is some clinical-experimental and a host of clinical-observational evidence to indicate that symbolic transformations are quite customary in human behavior, particularly where unacceptable or antisocial impulses are involved.

In case histories, conversion symptoms and compulsions frequently present themselves dramatically as evidence of the fact that a given conflict or forbidden impulse may secure expression in a manner only indirectly related to the original impulse. Thus, it is generally accepted that hand-washing compulsions often represent symbolically the desire of the individual to cleanse himself of the consequences of masturbation. Likewise, in doll therapy, there is excellent evidence that the child is able to discharge against surrogate objects the same impulses and feelings that he has developed toward persons in the real world.

The extent to which this relationship between symbol and real-life counterpart is unequivocally demonstrated varies with the individual case. In many instances, however, the relationship is effectively demonstrated as the symbolic behavior can be shown to vary directly with changes in the relationship between the subject and the referent of the symbol. Thus, the aggression directed against the father-doll may show a high inverse relationship to the extent to which the child is able to express such impulses against the original stimulus object—the father. Similarly, a symptom may disappear when the impulse or conflict that it is presumed to represent is dissolved by therapeutic procedure or through changes in the balance of environmental determinants. Levy has described a number of clinical cases involving relatively convincing demonstration of symbolic representation of sibling rivalry, fear of castration, defecation, and so on.³⁶ Tomkins reports the study of a single case under varying degrees of drunkenness. He found that, in certain cases, the hypothesized referent of the symbol

³⁵ R. A. McCleary and R. S. Lazarus, "Autonomic Discrimination without Awareness: An Interim Report," *Journal of Personality*, Vol. 18 (December, 1949), pp. 171-79.

³⁶ D. M. Levy, "Projective Techniques in Clinical Practice," *American Journal of Orthopsychiatry*, Vol. 19 (January, 1949), pp. 140-44.

was more and more directly represented as the amount of alcohol consumed increased.³⁷

Several investigators have attempted to combine the subject matter of clinical psychology with attempts at empirical control in the study of symbolic representation. An interesting study by Farber and Fisher attempted to demonstrate the process of symbolic transformation by means of using hypnosis and asking subjects to create and interpret dreams.³⁸ Although the study is not well controlled, if we accept the naïveté of the subjects, the evidence is impressive, for some of the subjects, that not only is symbolic representation an observed form of behavior, but also that the process of assigning symbols in large part corresponds with the rules established by Freud from his examination of dream protocols. Consistently, Krout, in a better-controlled investigation, studied interpretations based upon the assumption that individuals responded to line drawings, representative of basic experiences or objects, as if they were the symbol referent. She found that these interpretations could be verified through the inspection of independently collected validation data. In other words, the subjects did appear to respond to the line drawings in a manner consistent with the nature of the object that the Freudian would view as lying behind this symbol.³⁹

Franck and Rosen have reported two studies in which there is some evidence that the response of the individual to male and female symbols varies with his sex and relative maturity.⁴⁰ Klein, employing the method of hypnotically inducing dreams, was able to show that the dreamer characteristically transformed or disguised external stimuli that were incorporated into the dream.⁴¹

In general, then, controlled empirical evidence supplements clinical observation to imply the existence of tendencies for human subjects to represent dispositions and conflicts symbolically. The rules for de-

³⁷ S. S. Tomkins, *op. cit.*

³⁸ L. H. Farber and C. Fisher, "An Experimental Approach to Dream Psychology through the Use of Hypnosis," *Psychoanalytic Quarterly*, Vol. 12 (April, 1943), pp. 202-16.

³⁹ J. Krout, "Symbol Elaboration Test: The Reliability and Validity of a New Projective Technique," *Psychological Monographs*, Vol. 64 (1950), No. 4, Whole No. 310.

⁴⁰ K. Franck, "Preferences for Sex Symbols and Their Personality Correlates," *Genetic Psychology Monographs*, Vol. 33 (1946), pp. 73-123; K. Franck and E. Rosen, "A Projective Test of Masculinity-Femininity," *Journal of Consulting Psychology*, Vol. 13 (August, 1949), pp. 247-56.

⁴¹ D. B. Klein, "The Experimental Production of Dreams during Hypnosis," *University of Texas Bulletin*, No. 3009 (1930).

termining the symbol-referent relations are as yet imperfectly understood, although there is evidence indicating that the Freudian view of the symbolic process possesses some utility.

ARE THE STORIES THAT THE SUBJECT CREATES OF UNEQUAL IMPORTANCE
AS DIAGNOSTIC OF HIS IMPULSES AND CONFLICTS?

It is possible to assume that any behavioral information, no matter how scanty, contains in it the necessary elements to permit a complete understanding of the individual in question. This derivative of strict Freudian determinism is still defended by some investigators. It implies a theoretical model that is able to incorporate all possible empirical relationships and predict these relationships so precisely that, once a single value has been inserted into the closed theoretical system, every other construct in the system can be assigned a value. In view of the present state of psychological theory, this assumption seems an exceedingly unwise one. Nor does it appear to guide the actual behavior of most clinicians. The tendency to view different stories as possessing different degrees of psychological significance is rather clearly revealed by the fact that, not infrequently, even with twenty stories, the investigator will feel he has not sufficient material to make a diagnosis or come to any understanding of the dynamics of the individual in question. Apparently, just as some sets of twenty stories are more difficult to interpret and are less revealing, so also some individual stories are less rewarding than others.

Although presumably implicitly present in the minds of most interpreters, the assumption is seldom mentioned, and there has been little systematic attempt to state any criteria by means of which the more important stories are to be separated from the less important. Rapaport, who represents an exception to this generalization, has suggested that the more important stories may be distinguished in terms of certain "formal variables." He cites as illustration of these distinguishing characteristics: the consistency of the story with other stories of this individual, its consistency with stories told by other persons to this picture, the faithfulness with which the subject followed instructions in telling this story, and, finally, the extent to which the individual "has perceived and apperceived the picture adequately in all its parts."⁴²

Indirect evidence bearing on this question is provided by the common observation that stories derived from particular cards frequently

⁴² D. Rapaport, "The Clinical Application of the Thematic Apperception Test," *Meninger Clinic Bulletin*, Vol. 7 (May, 1943), pp. 106-13.

have little information pertinent to given variables. Thus the individual interested in studying aggression may find that stories told in response to certain TAT cards will only rarely provide pertinent evidence. Not only do the stories told by the subject vary in significance but, for given purposes, the stories characteristically evoked by different cards vary in significance. Eron has shown that both incidence of themes and the emotional tone of stories differ significantly among the various TAT cards.⁴³

Although practical considerations and clinical experience indicate the necessity of an assumption of "crucial stories," there seems to be little empirical evidence that is directly pertinent.

ARE THEMES OR STORY ELEMENTS THAT APPEAR TO ARISE DIRECTLY
OUT OF THE STIMULUS MATERIAL LESS SIGNIFICANT THAN THOSE
THAT ARE NOT DIRECTLY DETERMINED BY THE STIMULUS
MATERIAL?

The extent to which the assumption that story elements not directly determined by stimulus material are more significant than those that are is embraced by projective testers is made clear in the general emphasis upon the importance of bizarre responses, sex reversals, and so on. Rotter gives it explicit stress in his interpretive discussion of the importance of "unusualness" of response in the Thematic Apperception Test.⁴⁴ In the scoring of sentence completion responses, Rotter has reported finding special diagnostic significance associated with "twist" or "reversal" endings to incomplete sentences.⁴⁵ Weisskopf has emphasized the assumption heavily in her suggestion that one method of estimating the diagnostic efficiency of projective stimulus material is through the use of a "transcendence index." This index is derived from the descriptive comments of the subject in responding to the stimulus material that go beyond "pure description."⁴⁶

The first difficulty encountered in attempting to appraise this assumption is the question of how to go about measuring the degree to

⁴³ L. D. Eron, "A Normative Study of the Thematic Apperception Test," *Psychological Monographs*, Vol. 64 (1950), No. 9, Whole No. 315; L. D. Eron, D. Terry, and R. Callahan, "The Use of Rating Scales for Emotional Tone of TAT Stories," *Journal of Consulting Psychology*, Vol. 14 (December, 1950), pp. 473-78.

⁴⁴ J. B. Rotter, "Thematic Apperception Tests: Suggestions for Administration and Interpretation," *Journal of Personality*, Vol. 15 (September, 1946), pp. 70-92.

⁴⁵ J. B. Rotter, J. E. Rafferty, and E. Schachtitz, "Validation of the Rotter Incomplete Sentences Blank for College Screening," *Journal of Consulting Psychology*, Vol. 13 (October, 1949), pp. 348-56.

⁴⁶ E. A. Weisskopf, "A Transcendence Index as a Proposed Measure in the TAT," *Journal of Psychology*, Vol. 29 (April, 1950), pp. 379-90.

which a given response is determined by the stimulus material. This could be determined by estimating the structural similarities between the response and the stimulus. Or, one might develop empirical norms representing the common or usual response elements for each stimulus. Rosenzweig has chosen the second alternative in his attempt to establish "apperceptive norms." The function of these norms, in large part, is to enable the interpreter to differentiate that which derives naturally from the stimulus card from that which is projective or personally determined.⁴⁷ The norms reported by Eron could be used in similar fashion.⁴⁸ Presumably, most clinicians implicitly employ both of these approaches. On the one hand they note responses which are an obvious denial of the observable elements in the stimulus material, and on the other they build certain expectations in regard to what a "normal" story is to each stimulus, and they tend to place special interpretive emphasis upon departures from these norms.

Some projective testers probably object to this assumption on the grounds that "any" response, no matter how normal or directly derived from the stimulus material, may have significance in projective protocols. They might point to this attribute as one of the chief differentiae between mental testing (normative) and projective testing (idiosyncratic). It is certainly true that the extent to which a response deviates from the norm is much less a matter of concern to the interpreter of projective protocols than it is to the "mental tester." The fact remains, however, that if a particular response is given by 100 per cent of the persons taking the test, this response cannot possess diagnostic significance unless it is combined with some other response element that is less frequently encountered. In the latter case, it is clear that the second or less-frequent response is determining the interpretation, not the "normal" response. Nevertheless, it must be remembered that even a response given by a very high percentage of the respondents may have broad significance as in the case of a yes-no response to a specific question. In the limiting case, a response may be given by 99 per cent of the respondents and still be an important diagnostic sign, especially in those few cases where it is not encountered.

Limited empirical justification for the view that significance and "stimulus-boundedness" are negatively related is provided in a study by McClelland, Burney, and Roby. They observed that introduction

⁴⁷ S. Rosenzweig, "Apperceptive Norms for the Thematic Apperception Test: I. The Problem of Norms in Projective Methods," *Journal of Personality*, Vol. 17 (June, 1949), pp. 475-82; S. Rosenzweig and E. E. Fleming, "Apperceptive Norms for the Thematic Apperception Test: II. An Empirical Investigation," *Journal of Personality*, Vol. 17 (June, 1949), pp. 483-503.

⁴⁸ Eron, *op. cit.*

by the storyteller of an affiliated person into TAT stories where this person was not present in the picture was related to an experience the subjects had just undergone. A count of affiliated persons without consideration of whether or not they were in the picture did not reveal any relationship with the experimental treatment.⁴⁹

Supporting this assumption is the fact that most perception researchers seem to accept the generalization that the more ambiguous the material, the easier it is to observe the operation of directive or motivational factors. Thus, they suggest that in those cases where the response bears relatively little relation to the stimulus material we are especially apt to observe dynamic or motivational factors. Closely related also is Bruner and Postman's suggestion that the perceptual process may be represented as involving a relationship between the predetermining tendencies of the individual on the one hand, and information or stimulus constraints presented by the environment on the other. They imply that, the stronger the determining tendency, the less environmental supports are needed to produce a related percept and the more contradictory information will be needed to produce a percept in opposition to the predetermining tendency.⁵⁰ The above assumption is similar to Bruner and Postman's view, since it proposes that the stronger the stimulus constraints or supports, the less we know about the predetermining tendencies in the individual that have produced a response consistent with these constraints. If the stimulus material is not consistent with the report of the subject, however, presumably the predetermining tendency was sufficiently strong to produce an appropriate response, even though the stimulus did not call for it.

Thus, the assumption that stimulus-bound responses are less diagnostic than responses that do not depend so heavily upon stimulus constraints appears to fit well with available empirical data.

ARE RECURRENT THEMES IN A SERIES OF STORIES PARTICULARLY LIKELY TO MIRROR THE IMPULSES AND CONFLICTS OF THE STORYTELLER?

An assumption frequently stated, but for which there is a minimum of pertinent empirical evidence, is that recurrent themes are likely to

⁴⁹ D. C. McClelland, R. C. Burney, and T. B. Roby, "The Effect of Anxiety on Imagination," paper read at the 1950 meeting of the Eastern Psychological Association.

⁵⁰ J. S. Bruner, "Personality Dynamics and the Process of Perceiving," in R. R. Blake and G. V. Ramsey (eds.), *Perception, an Approach to Personality* (New York: Ronald Press, 1951), pp. 121-47; L. Postman, "Toward a General Theory of Cognition," in J. H. Rohrer and M. Sherif (eds.), *Social Psychology at the Crossroads* (New York: Harper and Brothers, 1951), pp. 242-72.

mirror the impulses and conflicts of the storyteller. There are heavy rational considerations, however, favoring such a view. The presence of the same theme, even when the stimulus situation has been thoroughly altered, implies strongly that there are impelling forces within the individual creating these themes rather than their being the inevitable outcome of stimulus constraints. The same theme following a change in stimulus material suggests that the response is not tied directly to, or produced by, a single specific stimulus. Thus, we may infer that the response tendency possesses generality and is more likely to be related to motivational factors than the response linked to a single stimulus.

Even if we accept the view that in all instances the stimulus is evoking the response at question, we know that one measure of the strength of a drive is the extent to which stimulus generalization or displacement will occur. Thus, if the individual can equate a large number of stimuli in order to make this response, we may infer that the instigation lying behind the response is quite strong. In addition, if interpretations are favored that incorporate a large amount of material, those based upon recurrent themes have a natural advantage in that the interpretation of one theme is automatically applicable to all of the other themes in the series. Finally, the presence of recurrent themes permits the investigator to sample or test the conditions under which these themes make their appearance.

DO STORIES REFLECT NOT ONLY ENDURING DISPOSITIONS BUT ALSO
MOMENTARY IMPULSES AND ARE THESE BOTH REFLECTED
IN THE SAME MANNER?

Most investigators are primarily interested in TAT as a means of measuring enduring dispositions, although, occasionally, especially in research, the measurement of situational or temporary instigations to behavior may be crucial.

There have been numerous studies showing the sensitivity of the instrument to temporary or situational determinants. Bellak showed that the number of aggressive words in TAT stories increased when the storyteller was rebuked for the low quality of the stories he told.⁵¹ Sanford, and Atkinson and McClelland have shown that TAT protocols vary with food deprivation,⁵² and McClelland, Clark, Roby, and Atkinson have shown story variation as a result of exposure to failure in

⁵¹ Bellak, *op. cit.*

⁵² Sanford, *op. cit.*; J. W. Atkinson and D. C. McClelland, "The Projective Expression of Needs: II. The Effect of Different Intensities of the Hunger Drive on Thematic Apperception," *Journal of Experimental Psychology*, Vol. 38 (December, 1948), pp. 643-58.

a test situation.⁵³ Rodnick and Klebanoff have shown that TAT stories vary with relative success in a level of aspiration test.⁵⁴ Lindzey demonstrated that extrapunitive behavior on the part of the hero in TAT protocols increased significantly following failure in a social situation.⁵⁵

In similar fashion, there are a number of studies that show the sensitivity of the stories to more enduring dispositions. It seems reasonable to accept individuals belonging to different psychiatric groups as differing in some enduring rather than situational attribute. Consequently, studies showing significant differences between groups separated on some diagnostic variable may be considered evidence of the test's sensitivity to enduring dispositions and conflicts. Balken and Masserman observed significant differences in the TAT performance of patients categorized as conversion hysterics, anxiety hysterics, and obsessive compulsives.⁵⁶ Renaud was able, with some difficulty, to distinguish between psychoneurotics, traumatic brain disorder cases, and brain disease cases on the basis of TAT protocols.⁵⁷ Cox and Sargent found differentiating signs in the TAT performance of "stable" and "disturbed" school children.⁵⁸ Working with mental hospital inmates, Harrison, in approximately 77 per cent of the cases, was able to identify accurately on the basis of TAT stories the psychiatric category in which the patient had been placed.⁵⁹ He reports similar results when "blind analysis" was employed, that is, when the tests were administered by a different person than the interpreter.⁶⁰

Further evidence of the sensitivity of the test to nonsituationally determined motivation is supplied by studies, such as Murray and Stein's, where a relatively high positive relationship was found between leadership ratings of ROTC candidates based on TAT per-

⁵³ D. C. McClelland, R. A. Clark, T. B. Roby, and J. W. Atkinson, "The Projective Expression of Needs: IV. The Effect of the Need for Achievement on Thematic Apperception," *Journal of Experimental Psychology*, Vol. 39 (April, 1949), pp. 242-55.

⁵⁴ E. H. Rodnick and S. G. Klebanoff, "Projective Reactions to Induced Frustrations as a Measure of Social Adjustment," *Psychological Bulletin*, Vol. 39 (July, 1942), p. 489 (abstract).

⁵⁵ Lindzey, *op. cit.*

⁵⁶ Balken and Masserman, *op. cit.*

⁵⁷ H. Renaud, "Group Differences in Fantasies: Head Injuries, Psychoneurotics, and Brain Diseases," *Journal of Psychology*, Vol. 21 (April, 1946), pp. 327-46.

⁵⁸ B. Cox and H. Sargent, "TAT Responses of Emotionally Disturbed and Emotionally Stable Children: Clinical Judgment versus Normative Data," *Journal of Projective Techniques*, Vol. 14 (March, 1950), pp. 61-74.

⁵⁹ R. Harrison, "Studies in the Use and Validity of the Thematic Apperception Test with Mentally Disordered Patients: II. A Quantitative Validity Study," *Character and Personality*, Vol. 9 (December, 1940), pp. 122-33.

⁶⁰ Harrison, "Studies in the Use and Validity of the Thematic Apperception Test with Mentally Disordered Patients: III. Validation by the Method of 'Blind Analysis,'" *Character and Personality*, Vol. 9 (December, 1940), pp. 134-38.

formance and leadership ratings independently executed by officers of the men in question.⁶¹ Likewise, Henry's study of the Navaho and Hopi suggests the ability of the test to discriminate these two groups,⁶² and further to supply psychological information concerning them that is consistent with information secured from extended observation or through the use of independent instruments. White showed that there was a significant relationship between TAT response and the dispositions leading to hypnotizability.⁶³ Harrison, in those cases where he was able to match descriptive statements based upon TAT responses with independent information derived from mental hospital case histories, was correct in 82.5 per cent of the instances.⁶⁴

The question of whether temporary and enduring tendencies are reflected in stories in exactly the same fashion is important for two reasons. First, the clinician is primarily concerned with the more enduring tendencies; therefore, it is desirable that he have some means of differentiating between these two classes of determinants. Second, some research, especially McClelland's, has implied that through studying the effect of situational factors upon TAT performance it is possible to arrive at means of interpreting TAT stories to reveal the operation of more enduring tendencies.⁶⁵ If the reflection of these two kinds of motivational factors should prove to be very similar or the same process, this would be unfortunate for the clinician, who would then be faced with the difficulty or impossibility of knowing whether a given tendency was temporarily instigated or whether this was a more permanent characteristic of the individual in question. On the other hand, such similarity in process would encourage experimental treatments as feasible means of approaching the task of specifying more exactly the means of inferring enduring tendencies.

McClelland has supplied some evidence that the enduring and temporary processes are reflected in the same manner.⁶⁶ He has demonstrated that the differences in subjects' TAT responses following experimental induction of a motivational state (threatening test situation) is related to other measures of the individuals' behavior, such

⁶¹ H. A. Murray and M. I. Stein, "Note on the Selection of Combat Officers," *Psychosomatic Medicine Monographs*, Vol. 5 (1943), pp. 386-91.

⁶² Henry, *op. cit.*

⁶³ R. W. White, "Prediction of Hypnotic Susceptibility from a Knowledge of Subjects' Attitudes," *Journal of Psychology*, Vol. 3 (1937), pp. 265-77.

⁶⁴ Harrison, "Studies in the Use and Validity of the Thematic Apperception Test with Mentally Disordered Patients: II. A Quantitative Validity Study," *op. cit.*

⁶⁵ Atkinson and McClelland, *op. cit.*; McClelland, Clark, Roby, and Atkinson, *op. cit.*

⁶⁶ D. C. McClelland, "Measuring Motivation in Phantasy: The Achievement Motive," in H. Guetzkow (ed.), *Groups, Leadership and Men: Research in Human Relations* (Pittsburgh: Carnegie Press, 1951), pp. 191-205.

as academic performance. This implies that the response of the subject to the immediate situation is related to his more permanent patterns of response.

McClelland and Liberman derived a system for scoring need achievement in TAT responses that was based upon the differences produced in TAT protocols following experimentally induced failure in a situation related to achievement.⁶⁷ In addition, they demonstrated that achievement, as measured by this scoring technique, was related to performance in an anagrams test and also related to the speed with which certain kinds of achievement-related words could be recognized when exposed tachistoscopically. Thus, a measure of need achievement derived from a temporary instigation appeared to be related to the initial instigating situation. While this evidence is not compelling, it does provide some support for the notion that temporary instigations affect TAT stories in a manner consistent with the way in which more enduring dispositions influence stories.

We appear to have excellent empirical evidence indicating that stories are responsive to both situational and enduring motivational factors. There is no conclusive evidence demonstrating the similarity or dissimilarity of the process whereby these two classes of determinants secure expression in the stories.

DO THE STORIES REFLECT EVENTS FROM THE PAST OF THE SUBJECT THAT HE HAS NOT HIMSELF ACTIVELY EXPERIENCED? ARE THESE EVENTS DIAGNOSTIC OF THE INDIVIDUAL'S DISPOSITIONS AND CONFLICTS?

Inquiry following the customary administration of the TAT has verified the hypothesis that individuals do incorporate material taken directly from scenes that they have witnessed or from movies or books to which they have been exposed. Murray has reported this.⁶⁸ Further, any clinician who has worked at all extensively with the TAT has inevitably many cases in his own experience of stories drawn directly from the world of the novel or drama, frequently with accompanying remarks indicating explicitly that this was the case. This is directly consistent with Freud's early dictum that each dream incorporates something from the events of the preceding day.⁶⁹

Given the influence of these nonparticipated events, the question

⁶⁷ D. C. McClelland and A. M. Liberman, "The Effect of Need for Achievement on Recognition of Need-Related Words," *Journal of Personality*, Vol. 18 (December, 1949), pp. 236-51.

⁶⁸ Murray, *Thematic Apperception Test Manual*, *op. cit.*

⁶⁹ S. Freud, "The Interpretation of Dreams," in A. A. Brill (ed.), *The Basic Writings of Sigmund Freud* (New York: Modern Library, 1938), pp. 181-549.

then becomes one of whether motivational factors affect or are revealed in the recall of such experiences. The selective function of memory and the importance of motivational determinants in the memory process have been recognized for some time, certainly since the appearance of Bartlett's treatise on memory.⁷⁰ A number of studies have investigated, under reasonably well-controlled conditions, memory as a function of such motivational variables as political ideology (Edwards,⁷¹ Levine and Murphy⁷²), value as measured by the *Study of Values* (McGinnies and Bowles⁷³), sex membership (Clark⁷⁴), mental set (Carmichael, Hogan and Walter⁷⁵), attitude (Postman and Murphy⁷⁶), and punishment (McGranahan⁷⁷). Although there are many other factors, such as primacy, recency, and vividness, known to influence recall, it seems reasonably well established that motivational factors do serve as one important class of determinants of memory.

If we accept these studies as evidence of the extent to which memory is influenced by motivational factors, it seems reasonable that the individual, in the process of recalling past events or experiences, will reveal or expose important aspects of his own motivational state. Consequently, the assumption that the particular events remembered by the subject are diagnostic of his dispositions and conflicts appears to be supported by available empirical data.

DO STORIES REFLECT GROUP-MEMBERSHIP AND CULTURAL OR SOCIAL DETERMINANTS AS WELL AS PERSONAL OR INDIVIDUAL DETERMINANTS?

The assumption that stories reflect group, cultural, or social determinants simply implies consistent differences between the fantasy

⁷⁰ F. C. Bartlett, *Remembering* (London: Cambridge University Press, 1950).

⁷¹ A. L. Edwards, "Political Frames of Reference as a Factor Influencing Recognition," *Journal of Abnormal and Social Psychology*, Vol. 36 (January, 1941), pp. 34-50.

⁷² J. M. Levine and G. Murphy, "The Learning and Forgetting of Controversial Material," *Journal of Abnormal and Social Psychology*, Vol. 38 (October, 1943), pp. 507-17.

⁷³ E. McGinnies and W. Bowles, "Personal Values as Determinants of Perceptual Fixation," *Journal of Personality*, Vol. 18 (December, 1949), pp. 224-35.

⁷⁴ K. B. Clark, "Some Factors Influencing the Remembering of Prose Material," *Archives of Psychology*, New York, No. 253 (July, 1940).

⁷⁵ L. Carmichael, H. P. Hogan, and A. A. Walter, "An Experimental Study of the Effect of Language on the Reproduction of Visually Perceived Form," *Journal of Experimental Psychology*, Vol. 15 (February, 1932), pp. 73-86.

⁷⁶ L. Postman and G. Murphy, "The Factor of Attitude in Associative Memory," *Journal of Experimental Psychology*, Vol. 33 (September, 1943), pp. 228-38.

⁷⁷ D. V. McGranahan, "A Critical and Experimental Study of Repression," *Journal of Abnormal and Social Psychology*, Vol. 35 (April, 1940), pp. 212-25.

productions of individuals who belong to, or have been socialized in, different social groups. Thus, a certain amount of the variation in any TAT production can be accounted for by the fact that the individual has grown up in a given milieu or social role. The importance of the assumption derives from the fact that overlooking this kind of variation introduces a serious source of error in the interpretation of imaginative protocols from members of more than one social group.

Although little effort has been made to explore the variations in fantasy productions between many of the important groups of our own society, it is quite widely accepted or expected that these differences exist. Even such an important cleavage as that between male and female has been little explored so far as TAT behavior is concerned, while such variables as socioeconomic status, occupational role, and ethnic-group membership have also been of slight interest to most investigators. Rosenzweig and Fleming have reported differences between a roughly equated group of men and women on a number of specific aspects of TAT response.⁷⁸ An investigation by Riess, Schwartz, and Cottingham,⁷⁹ designed as a critical appraisal of Thompson's modification of the TAT,⁸⁰ led to the observation of certain relatively slight variations in story length as a function of geographic residence and Negro-white group membership.

Henry, in his investigation of Navaho and Hopi children, found that inferences based upon his adaptation of the TAT related to independently secured information concerning the children and also that there were systematic differences between the Navaho and Hopi fantasy productions. In addition, he reports differences between Navaho subjects who were members of different subgroups in this society.⁸¹

Although there is a paucity of empirical evidence demonstrating differences in fantasy production between various sociocultural groups, what evidence is available appears to support this assumption.

ARE IMPULSES AND CONFLICTS INFERRED FROM STORIES NOT ALWAYS REFLECTED DIRECTLY IN OVERT BEHAVIOR AND CONSCIOUSNESS?

Projective testers vary considerably in the manner in which they emphasize the assumption that impulses and conflicts inferred from

⁷⁸ Rosenzweig and Fleming, *op. cit.*

⁷⁹ B. F. Riess, E. K. Schwartz, and A. Cottingham, "An Experimental Critique of Assumptions Underlying the Negro Version of the TAT," *Journal of Abnormal and Social Psychology*, Vol. 45 (October, 1950), pp. 700-709.

⁸⁰ C. E. Thompson, "The Thompson Modification of the Thematic Apperception Test," *Rorschach Research Exchange*, Vol. 13 (December, 1949), pp. 469-78.

⁸¹ Henry, *op. cit.*

stories are reflected directly in overt behavior. Some see a very intimate relation between imaginative behavior and overt behavior. Thus, Piotrowski introduces nine rules designed to permit the translation of fantasies into overt behavior with the following statement:

The rules proposed in this article are a new attempt to solve the problem of the relationship between the TAT and overt behavior. Since the TAT is mainly an exercise in creative imagination, it should reflect the patient's ideas and drives regardless of whether or not they find a direct expression in overt behavior. Thus, parts of the TAT always reflect the overt behavior of the subject while other parts reflect ideas which are not as directly manifested in overt actions. If this be so, we need a rule by means of which we could differentiate these two parts of the TAT. The rules presented below have been formulated largely for the purpose of meeting that need.⁸²

Others are more cautious in stating their views of the relation between story behavior and overt behavior. Murray, in his introduction to the TAT, suggests:

It may be stated, as a rough generalization, that the content of a set of TAT stories represents second level, covert . . . personality, not first level, overt or public . . . personality. There are plenty of ways of discovering the most typical trends; the TAT is one of the few methods available today for the disclosure of covert tendencies. The best understanding of the total structure of personality is obtained when the psychologist considers the characteristics of manifest behavior in conjunction with the TAT findings . . .⁸³

In similar vein, Korner states:

Instead of deploring the fact that fantasy and reality behavior do not necessarily correspond, as we currently seem to be doing, we can use projective techniques as a shortcut to a person's fantasy and ideational life, which then can be compared and examined in the light of his present and past actual behavior patterns.⁸⁴

Although the relationship between covert and overt is assessed differently by various investigators, all seem to agree that the relationship is not perfect—fantasy behavior does not exactly mirror overt behavior. This omnipresent assumption can serve one of two functions, depending upon the orientation of the investigator. It *can* serve simply as a convenient means of avoiding the necessity of ever being wrong. Thus, whenever inferences based on story protocols fail to relate to appropriate independent measures or observations, the clinician may

⁸² Piotrowski, *op. cit.*, p. 105.

⁸³ Murray, *Thematic Apperception Test Manual*, *op. cit.*, p. 16.

⁸⁴ A. F. Korner, "Theoretical Considerations Concerning the Scope and Limitations of Projective Techniques," *Journal of Abnormal and Social Psychology*, Vol. 45 (October, 1950), p. 627.

simply point to the above assumption and add that only the naïve would expect always to observe linear relationships between imaginal and overt or conscious behavior. On the other hand, the investigator can use this assumption as a signpost pointing to one of the most important and difficult empirical problems facing the projective tester. This problem is the determination of the conditions under which inferences based on projective material directly relate to overt behavior and the conditions for the reverse.

It is possible to defend the position that projective techniques should not be expected to provide statements concerning overt behavior. Such a view implies that the techniques will always be used only as an "imaginal supplement" to an otherwise adequate description of the individual. Thus, given a person who "behaves" in a particular way, examination of his fantasy productions may permit us to make consistent interpretations or explain behavior hitherto unaccountable. Certainly this represents an important function of these techniques. Equally certain is the fact that this is not the only circumstance under which these instruments are used. They *are* used as means of inferring overt behavior tendencies, and presumably with more adequate rules of transformation they would be so used much more widely.

Investigations by Sanford and co-workers and Symonds, which unfortunately from the point of view of sampling were both based on adolescent populations, demonstrate clearly that in some cases instead of the impulses inferred from TAT records being reflected in behavior, their converse or opposite appear in behavior.⁸⁵ This observation leads to the question of whether impulses that secure release in overt behavior may not need to be expressed in fantasy productions. The many positive relationships, however, between imaginal impulses and overt behavior in these studies and others make it clear that it is not an either-or proposition and that the statement of the actual conditions under which the impulse is revealed or concealed must be complex.

Sanford's group studied the relationship in a group of school children between fantasy ratings derived from the TAT and overt behavior ratings provided by teachers who had observed the children. They found an average correlation of $+ .11$ between the two sets of ratings, indicating clearly that the fantasy ratings alone were not good

⁸⁵ R. N. Sanford, M. M. Adkins, R. B. Miller, *et al.*, "Physique, Personality and Scholarship: A Cooperative Study of School Children," *Monographs of the Society for Research in Child Development*, Vol. 8 (1943), No. 1; P. M. Symonds, *Adolescent Fantasy: An Investigation of the Picture-Story Method of Personality Study* (New York: Columbia University Press, 1949).

predictors of overt behavior. There were striking differences, however, between the different variables used in the extent to which fantasy and behavior corresponded. For some needs there was a relatively high positive relationship, while for others there was a significant negative relationship between the overt and covert. In accounting for these findings, Sanford and his co-workers suggested that those tendencies which were negatively sanctioned or prohibited would be high in fantasy and low in overt behavior, while those tendencies which were encouraged by society and for which the individual could secure complete overt expression would be high in behavior but low in fantasy. High ratings would be secured in both fantasy and overt behavior for those tendencies that society encouraged but did not permit complete freedom of expression in, for example, achievement and dominance.

Murray has suggested that tendencies not inhibited by cultural sanctions are apt to be highly correlated in their fantasy and overt expression. He reports a positive correlation of over .40 between fantasy and overt behavior for a group of college men on the following variables: abasement, creation, dominance, exposition, nurturance, passivity, rejection, and dejection. Negative correlations are reported between fantasy and overt for sex, and no correlation between the two forms of expression and achievement.⁸⁶ Korner attempted to relate hostility as observed in a play situation with ratings of hostility in interpersonal relations with other children. She found no general relationship between the two sets of variables. Half of the children high on hostility in play situations were likewise high on hostility manifested in their dealings with other children. The remaining half who were high on hostility manifested in play situations were low on the second set of ratings. The investigator concluded that it was impossible to predict from the one situation to the other.⁸⁷

Symonds related the fantasy themes of 40 adolescent boys and girls to adjustment ratings and teachers' ratings of behavioral characteristics. He concluded that the relationship between these two sets of variables was "insignificant and negligible."⁸⁸

Further evidence of the lack of a perfect relationship between fantasy and overt behavior is provided in those cases where an individual of known characteristics fails to reveal salient aspects of himself in his TAT constructions. Tomkins reports the case of an individual who had

⁸⁶ Murray, *Thematic Apperception Test Manual*, op. cit.

⁸⁷ A. F. Korner, *Some Aspects of Hostility in Young Children* (New York: Grune and Stratton, 1949).

⁸⁸ Symonds, op. cit.

a persistent spontaneous fantasy which included as an important theme a homosexual seduction. The TAT responses of this individual gave no sign of homosexual tendencies. In accounting for this and similar cases, Tomkins suggests that the important variables are the awareness of the subject of the impulse or tendency at question and the extent to which the tendency is condoned or accepted by society. If the impulse is known and unaccepted by society, the individual will prevent its appearance in the stories he tells. If he is unaware of the tendency, it will appear in his fantasy constructions even if it is negatively sanctioned by society.⁸⁹ Bellak reports several similar instances where TAT performance fails to reveal central aspects of the individual.⁹⁰

Relatively little has been done in the attempt to discover and formulate signs in the stories themselves that would provide evidence concerning the probability of overt expression. There is some evidence that, as Tomkins proposes, the "psychological distance" maintained by the storyteller toward the impulse or disposition in question may be an important condition relating to the degree of overt translation.⁹¹

Available empirical evidence clearly indicates that the assumed imperfect correlation between fantasied and overt behavior is warranted. At present, however, we are far from an adequate formulation of the signs or cues that might permit specification from fantasy protocols alone of the behavioral tendencies that will secure overt expression as opposed to those that will not.

RESEARCH IMPLICATIONS

Almost all of these assumptions point to further research that would be useful in clarifying their status. Perhaps more important than research aimed at further demonstrating the warranty of these same assumptions is research that attempts to provide a more exact statement of the conditions under which the assumptions are applicable and the way in which they can be related to empirical data. For example: what are the means by which the important story in a series can be determined? How can we determine whether or not a given fantasy impulse will receive overt expression? In what way do we determine whether or not a given response has been determined by the stimulus material? What are the circumstances under which symbolic transformations must be engaged in? How do we determine the empirical

⁸⁹ Tomkins, "The Present Status of the Thematic Apperception Test," *op. cit.*

⁹⁰ L. Bellak, "Thematic Apperception: Failures and the Defenses," *Transactions of the New York Academy of Science*, Vol. 12 (February, 1950), pp. 122-26.

⁹¹ Tomkins, *The Thematic Apperception Test*, *op. cit.*

referent of a given symbol? Answers to these and a host of related questions are necessary before we can hope to provide the TAT user with an explicit, repeatable set of operations for inferring motivational states.

In addition to problems connected with the interpretive assumptions and the more specific questions implied in the above paragraph, there is also the matter of formulating explicitly a method of scoring TAT protocols that is practical, intersubjective, and able to embrace a reasonable number of the behavioral variables in common use. To a large extent standardization and specific interpretive rules must wait until some agreement has been reached by most TAT users as to the major aspects of TAT response that will be focused upon in analysis.

15. THE ERROR-CHOICE TECHNIQUE*

Here is an ingenious new variant on projective techniques which as yet seems to have received little attention in marketing circles. Admittedly, the technique is a highly controversial one, and much remains to be done in establishing its reliability. The basic idea, however, of measuring attitudes, or preferences, through the use of multiple-choice questions containing nothing but wrong answers seems sufficiently intriguing to merit consideration at least from marketing researchers.

There can be little dispute with the growing restlessness over present methods of attitude measurement.¹ Impressed with the results of projective techniques used in personality diagnosis, social psychologists show signs of seeing possibilities in this method or, at least, in approaches to it. This article will be concerned with the exploratory development and application of a method which, while not deserving the term "projective," might at least be termed "indirect."

Since much of the difficulty with present methods of attitude measurement lies in the trouble authors have in deciding just what it is they are trying to measure, I would suggest that use be made of experimentally derived, rather than logically derived, concepts con-

* Adapted from an article by Kenneth R. Hammond, University of California, "Measuring Attitudes by Error-Choice: An Indirect Method," *Journal of Abnormal and Social Psychology*, Vol. 43 (1948), pp. 38-48.

¹ Q. McNemar, "Opinion-Attitude Methodology," *Psychological Bulletin*, Vol. 43 (July, 1946), pp. 289-374.

cerning attitudes and consider an attitude as a (nonprimary) source of energy, or an affective state, capable of producing error in perception and recall, recognizing that the nature of this source of energy or affective state is still only roughly delineated.

The trap of definition will be avoided, then, by being concerned only with the effect of attitude, and the problem of definition will be left to those who are willing to be concerned with it. The particular effect with which we shall be concerned here will be the systematic error in perception and recall. The technique to be developed strives to leave the subject no alternative save error, eliminating reality as a factor, thereby affording a measure of the constancy and direction of the error.

EXPERIMENTAL BACKGROUND

The fact that perception may be influenced or that systematic errors may be induced has long been established. Sherif has surveyed the literature² and noted the evidence concerning the constancy of the direction of error in experimental situations:

The literature is rife with data which support the formulations reached from our survey of general psychology; that perception and judgment are selective and operate within a referential framework, . . . once established these frames and points of reference serve as anchorages for perception and judgment.

Sherif has concerned himself with this problem in connection with the autokinetic effect produced by a stationary pinpoint of light in a dark room.³ His experience with this phenomenon leads him to this point of view:

The first stage of attitude formation—in the most complicated social situations as well as in a restricted laboratory experiment—is a perceptual stage. . . . [Moreover,] in many cases the objective set is dominant in the situation. There are cases, however, in which this objective determination is lacking, thus allowing internal factors such as attitudes, subjective norms, and values to play the dominant role in organization of the perceptual field.⁴

The "cases in which this objective determination is lacking" certainly occur in the field of social psychology.

Sherif's work is important not only because of his own experiments but also because he has brought to focus a great portion of data from

² M. Sherif, *The Psychology of Social Norms* (New York: Harper and Brothers, 1936); M. Sherif, "A Study of Some Social Factors in Perception," *Archives of Psychology*, No. 187 (July, 1935); M. Sherif and H. Cantril, "The Psychology of 'Attitudes,' Part II," *Psychological Review*, Vol. 53 (January, 1946), pp. 1-24.

³ Sherif, *The Psychology of Social Norms*, *op. cit.*

⁴ Sherif, "A Study of Some Social Factors in Perception," *op. cit.*

experimental psychology in such a manner as to provide a valuable tool for social psychology.⁵

For our purposes the importance of the many previous laboratory experiments is this: The field of social events and personalities presents complex and confusing stimuli to the observer. If we know that the perception of such stimuli may be distorted by social factors, we then have the possibility of utilizing this phenomenon.

As for selective forgetting, Edwards has surveyed the literature and concludes that experiences which harmonize with an existing frame of reference will not only be learned and remembered better than conflicting experiences but that experiences that are in opposition will be recast and thus assimilated more rapidly, and the chances are slight that conflicting experiences will cause serious reorganization. Edwards' own experiments support these conclusions.⁶

The significance of these prior experiments for our purpose is this: The error-choice method used here provokes the subject to draw upon his memory of events in order to decide which answer is "correct." Since the field from which he does draw is ambiguous and confused at best, we know from the above evidence that the subject will select those pseudo-facts from memory which fit his frame of reference or support his established premise. Recall under these circumstances can be demonstrated to be selective, that is, nonrepresentational, therefore erroneous.

The instrument for measuring the effect of attitude, then, will make use of these two dynamics: (1) the distortion of perception and (2) the selective recall of the previous perceptual experience.

A general statement of the problem can now be made as follows: We are attempting to measure the effect of attitude, herein considered

⁵ Others have contributed positive and interesting results which bear on the problems of social dynamics affecting perception. See H. Proshansky and G. Murphy, "The Effects of Reward and Punishment on Perception," *Journal of Psychology*, Vol. 13 (1942), pp. 295-305; R. Wallen, "Individual Estimates of Group Opinion," *Journal of Social Psychology*, Vol. 4 (1940), pp. 25-29; M. Zillig, "Einstellung und Aussage," *Zeitschrift der Psychologie*, Vol. 106 (1928), pp. 58-106; E. L. Thorndike, "A Constant Error in Psychological Ratings," *Journal of Applied Psychology*, Vol. 4 (March, 1920), pp. 25-29.

⁶ A. L. Edwards, "The Retention of Affective Experiences: A Criticism and Restatement of the Problem," *Psychological Review*, Vol. 49 (January, 1942), pp. 43-53; "Political Frames of Reference as a Factor Influencing Recognition," *Journal of Abnormal and Social Psychology*, Vol. 36 (January, 1941), pp. 34-50. Further studies may be found in: J. M. Levine and G. Murphy, "The Learning and Forgetting Controversial Material," *Journal of Abnormal and Social Psychology*, Vol. 38 (October, 1943), pp. 507-17; W. S. Watson and G. W. Hartmann, "The Rigidity of a Basic Attitudinal Frame," *Journal of Abnormal and Social Psychology*, Vol. 34 (July, 1939), pp. 314-35; R. Wallen, "Ego-Involvement as a Determinant of Selective Forgetting," *Journal of Abnormal and Social Psychology*, Vol. 37 (1942), pp. 20-39. For an extensive survey and bibliography, see Sherif and Cantril, *op. cit.*

to be a (nonprimary) source of energy, or affective state, capable of distorting perception and recall with reference to an unstructured, ambiguous world of social events, by measuring the constancy of the direction of the error into which the respondent will be forced.

QUESTIONNAIRE DEVELOPMENT

Three series of items were presented under the guise of an "information test."⁷ One series of 8 questions was made up with alternate answers equidistant from the truth in opposite directions. [Example: "Average weekly wage of the war worker in 1945 was (1) \$37, (2) \$57."] In this case the facts were determinable. The second series of 12 questions also offered alternative opposing answers. [Example: "Russia's removal of heavy industry from Austria was (1) legal, (2) illegal."] In this case, the facts were indeterminable. There was, then, a total of 20 of these two types of questions in each test. It was the responses to these items, hereafter referred to as "nonfactual" items, that were scored. The third series of questions consisted of 20 straight information items, which were interspersed among the 20 items mentioned above. These will be referred to as "factual" items and were introduced in order to aid in disguising the test.

One test concerned Russia and consisted of 20 factual and 20 nonfactual questions. A second test concerned Labor-Management relations and also consisted of 20 factual and 20 nonfactual questions.

The subjects were told that these were information tests, that they would probably not know the answers to all the questions, that when they were in doubt they should guess, and emphasis was placed on working as rapidly as possible.

In constructing an item, the principal requirement was to eliminate reality, the truth of the matter, as a factor and thus force the respondent into a choice of errors and still make the item sound like an information item. Reality was eliminated (1) by putting the answers equidistant from the truth in opposite directions and (2) by using

⁷ See T. M. Newcomb, "The Influence of Attitude Climate upon Some Determinants of Information," *Journal of Abnormal and Social Psychology*, Vol. 41 (July, 1946), pp. 291-302, who discovers the relationship of an information test to attitudes. He notes that a certain information test was difficult, thus more guessing was to be expected and "the direction of the guessing is altogether likely to be weighted toward the subject's attitude. If this reasoning is correct, the . . . test tends to become itself an attitude test."

See also G. Myrdal, *An American Dilemma* (New York: Harper, 1948): "In most cases, the indirect analysis of the valuation sphere, through the study of the deviations of beliefs from true knowledge, is likely to reach deeper than does the direct analysis. An individual continually tends to arrange his valuations so that they may be presented in an acceptable form. But in his beliefs concerning social reality—which are shaped to give the appearance of rational organization to his morals—he reveals himself."

questions where the truth is indeterminable and putting the answers to opposite extremes.

The original selection of the items was based principally on "hunch." Since these two controversial subjects, Russia and Labor-Management relations, have given rise to a maze of contradictory and confusing "facts," I could only guess which "facts" would provide items that would sound like information items and at the same time afford logical justifications for a pro and anti premise. The check on my selection of items was afforded by the following validation procedure.

For example, if a group of people whose ways of life are such that one would expect a bias in favor of Labor and Russia makes constant errors in a pro direction on the items, we then have a certain substantiation of the "hunch" that these items would provoke positive errors. If we go farther, and find that a group of people whose ways of life are such that one would expect a bias against Russia and Labor makes constant errors in an anti direction, then we have a double substantiation of the "hunch" concerning the items—they provoke constant errors in opposite directions for groups with opposing bias. This was chosen as the criterion for an item: Does it or does it not discriminate between groups with opposing bias, that is, does the item discriminate between "known groups"?

My hypothesis in connection with the instrument developed, then, was this: Given two groups of individuals whose ways of life indicate beyond all reasonable doubt that they have opposing frames of reference concerning Labor and Russia, the members of these groups will err systematically in opposite directions. In statistical terms this means that the means of these groups will be significantly different in the direction predicted, that is, the pro-Labor group will have a high mean score, the anti-Labor group will have a low mean score, and the same in connection with the Russia form.

VALIDATION GROUPS

The experimental validation groups consisted of: (1) a group whose ways of life were such that one would expect a bias in favor of Labor and Russia, consisting of 18 adults who are employed by a major labor organization in clerical and semiprofessional positions (hereafter referred to as "Union"), (2) a group whose ways of life were such that one would expect a bias against Labor and Russia. This latter group was made up of two businessmen's luncheon clubs. One club, hereafter referred to as "Bus 1," was made up of 23 middle-aged

businessmen whose income varied from \$10,000 a year up. They were, in general, employers. The second club, hereafter referred to as "Bus 2," was fostered by "Bus 1" and was composed of 19 younger businessmen in their thirties.

To sum up, the subjects consisted of a group of people who work for a labor organization and two groups of people gathered together on the basis of having common business interests.

RESULTS

The two filled-out questionnaires, Labor and Russia, each include 20 nonfactual items to be scored. *A priori* determined positive (or pro) systematic errors were given a score value of one on each item; negative errors, no value. Thus, a high score indicates a pro bias, or positive systematic error. A low score indicates the reverse.

TABLE 1
LABOR QUESTIONNAIRE

<i>Group</i>	<i>Mean</i>	<i>SD</i>	<i>N</i>
Bus 1	8.21	3.45	23
Bus 2	9.93	3.00	19
Union	18.1	.84	18

TABLE 2
RUSSIA QUESTIONNAIRE

<i>Group</i>	<i>Mean</i>	<i>SD</i>	<i>N</i>
Bus 1	6.52	4.47	23
Bus 2	7.26	3.89	19
Union	16.5	1.96	18

The labor questionnaire means and sigmas are presented in Table 1. The Russia questionnaire means and sigmas are presented in Table 2.

Fisher's test for significance of differences of means between small samples shows no significant difference between groups "Bus 1" and "Bus 2." It is interesting, however, that the questionnaire was able to elicit a difference in means in the direction anticipated between a more structured, older, business group and a less homogeneous, younger group. On the basis of no significant difference between these two

strata, however, they were combined and compared with the labor group. The difference in the means between these two groups (Bus_{com} and Union) is highly significant.

Concerning validation of items, see Figure 1 for a statistical and graphic presentation. Each item is plotted for the percentage of labor

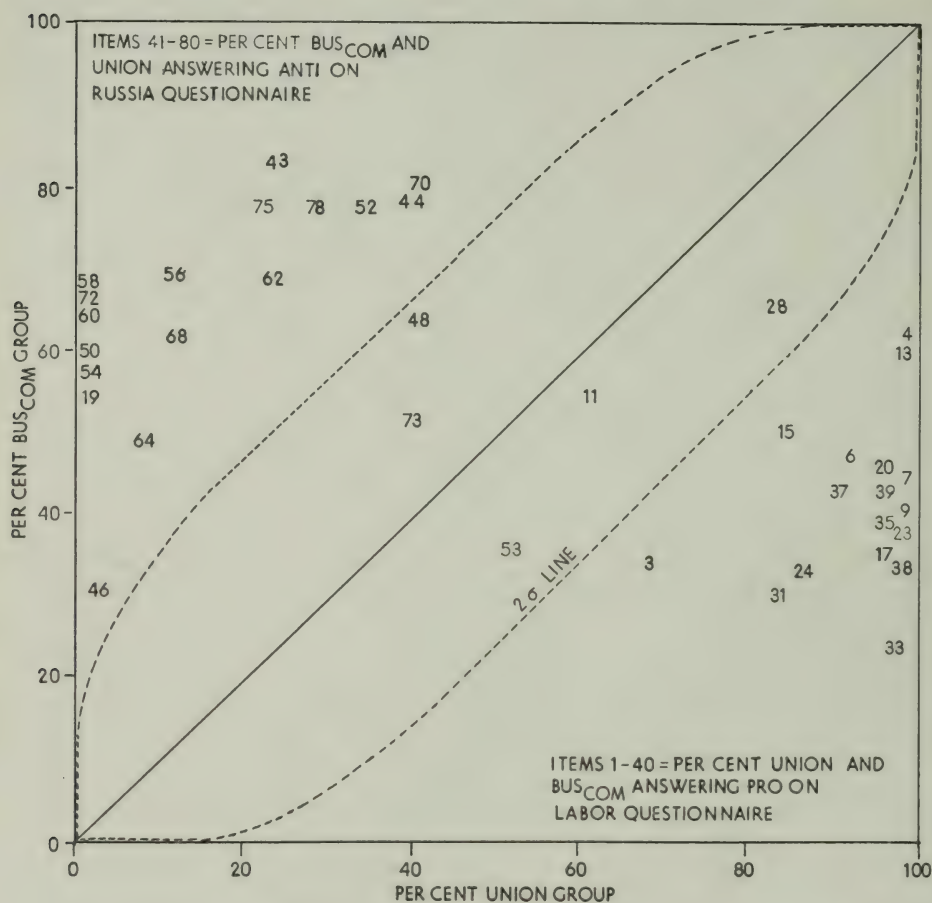


Fig. 1. Item validity: percentage of criterion groups answering "nonfactual" items the same way.

and business group subjects answering the item in the same way, that is, making an error in the same direction. For example, on Item 33 on the Labor questionnaire, 100 per cent of the Union group answered in a pro-Labor manner while only 23 per cent of Bus_{com} answered it in a pro-Labor manner. On Item 43 on the Russia questionnaire, 83 per cent of the Bus_{com} group answered in an anti-Russia manner, while only 21 per cent of the Union group answered in an anti-Russia manner. Those items falling outside the dotted line

show a difference in error-choice significant at the 5 per cent level.⁸

The ease with which high discriminatory values for items may be secured by this technique is impressive when one considers that selection of items was based principally on "hunch." It should be noted that this summary treatment produced only 5 poor items out of 40.

Reliability coefficients were obtained by the split-half method from a sample made up of 10 cases from the Union group, 40 cases from the Buscom group, and 10 college students chosen at random from a large sample used as an experimental group under circumstances to be discussed later. These cases were used in an effort to obtain a heterogeneous sample somewhat representative of the population. The reliability coefficient obtained on the Labor questionnaire, corrected by the Spearman-Brown formula, was .78. On the Russia questionnaire, the corrected reliability coefficient was .87.

POSSIBLE IMPROVEMENTS

Suggestions for improving the technique are as follows:

1. The original selection of items should be based on a definite objective. Validation is not the only goal in connection with this technique. It should also provide a definite clue as to the prevailing set of "factual" justifications which a sample is using. The suggestion is made that, rather than selecting items by "hunch," one should make use of some type of content, or symbol analysis, such as White's "value-analysis," in selecting items.⁹

2. The test should be carefully disguised. Present experience indicates that perhaps two neutral items to one "nonfactual" item is necessary.

⁸ The locus of the dotted 2 sigma line is plotted from points derived from a formula developed by George Kuznets and R. C. Tryon of the University of California. The formula for the case sigma equals 2 is

$$P_2 = \frac{(N_2 p_1 + 2) \pm 2 \sqrt{1 + p_1 q_1 \frac{N_2}{N_1} (N_1 + N_2 + 4)}}{N_1 + N_2 + 4}$$

where P_2 = point to be plotted on ordinate

p_1 = given point on abscissa

$q_1 = (1 - p_1)$

N_1 = Number of cases in group plotted on X axis

N_2 = Number of cases in group plotted on Y axis

This graphic method of determining significant differences is an analogue of that proposed by J. Zubin in "Note on a Graphic Method for Determining the Significance of the Difference between Group Frequencies," *Journal of Educational Psychology*, Vol. 27 (September, 1936), pp. 431-44, in which the procedure is a development of Zubin's C_1 equation.

⁹ R. K. White, "Value Analysis," *Journal of Social Psychology*, Vol. 19 (May, 1944), pp. 351-58.

3. It is possible to offer more than two answers. It seems probable that tests constructed with four error-choices to provide for "intensity" of error would prove useful for scaling items.

4. Not mentioned above was the fact that, for purposes to be developed later, each factual item was paired with a "nonfactual" item, thus affording a bona fide information test on precisely the same material on which error-choices were based. It is thus possible with this technique to correlate the information level of a group with their systematic error-choices. The factual items used to disguise the test therefore need not be dead weight; rather, they make the instrument more productive.

THIS TECHNIQUE IN RELATION TO TEST "SET"

Interest is increasingly becoming focused upon the problem of attitude test "set." It is the same problem that turned clinical psychologists from personality questionnaires to the use of projective techniques. The question is whether or not the subject is responding to the test items in terms of the test situation. Clinicians have learned that projective techniques have more often than not turned up dynamics that were rejected or denied by the subject in answer to direct questions. With this in mind, the question was raised of whether or not the error-choice technique would evoke responses different from those evoked by a direct method.

The problem was stated as follows: Will the error-choice technique evoke a different sort of response to the same item depending upon whether the items are presented as "information test" items or as "attitude test" items? Will the response be different under conditions of awareness from that under conditions of unawareness? That is, holding the item constant, will responses vary with a change in the respondent's "set" toward the test?

This problem was made subject to experiment in the following manner. The "information test" given to the validation groups was split into two forms. One form was that originally given to the validation groups, and was labeled "INFO." Form INFO again was given as an "information test." The other form contained the nonfactual items separated from and preceding the factual items, and was labeled "ATT-INFO."¹⁰ The first part of ATT-INFO, containing the nonfactual items, was preceded by instructions to the effect that this was an

¹⁰ Nonfactual items are the items upon which the error-choice scores were computed in connection with the validation groups. Factual items are the straight information items interspersed among the nonfactual items in order to aid in disguising the test.

"attitude test," that neither of the answers were correct, and that one should indicate his "attitude, his feeling" about the question, and instructions were given how to mark a pro or anti answer. The part containing the straight information items was preceded by instructions to the effect that this was an information test.¹¹

These two forms were administered to a sample consisting of students in an elementary psychology class; 97 subjects received form INFO, 47 received form ATT-INFO.

Results. The response on each item under both forms is expressed graphically and statistically in Figure 2. By plotting the percentage of the subjects who answered a given item in the same way on both INFO and ATT-INFO, as in Figure 2, we can discover whether subjects responded to the error-choice technique (INFO) in the same manner in which they responded to the questionnaire which expressly calls for attitudes (ATT-INFO).

On the Labor questionnaire we find that the items tend to fall along the equal percentage line, and few fall outside the dotted (2-sigma) line. In other words, about the same percentage in each group answered the same item in the same way. The items are obviously positively correlated. It should be noted, however, that about one third of the items do fall outside the 2-sigma line; that is, a change did occur in about one third of the items significant at the 5 per cent level.

On the Russia questionnaire, Figure 3, we find that more items are scattered away from the equal percentage line and that more of them show significant differences. In this case, the items appear somewhat negatively correlated. Responses to the same item tended to be different depending upon whether they were presented as part of an "attitude test" or as part of an "information test." In this instance, then, responses varied with awareness, with the "set" the respondents took toward the test situation.

Other concomitant effects of presenting the same item under a different test situation may be briefly stated: (1) The mean scores did not vary significantly; (2) variability tends to increase on that error-choice form which produces the greater shift in responses; (3) reliability, although low (.33 on the Labor form, .51 on the Russia form, both corrected by the Spearman-Brown formula), also tends to

¹¹ The order of the information questions and attitude questions was reversed on a third form given to the subjects for control purposes (INFO-ATT). No significant difference occurred between these forms, i.e., it made no difference in the mean score whether attitude items came first or last. The ATT-INFO form, however, was used for comparative purposes since it is not preceded by a test, even though this does not appear to be a disturbing factor.

increase on the form which produces the greater shift in responses.

These low reliability coefficients, obtained under circumstances which ordinarily provide very high reliability coefficients, are provocative and lead me to offer this resultant hypothesis: Reliability is a function of the control the subjects have over reproducing attitudes which they deem

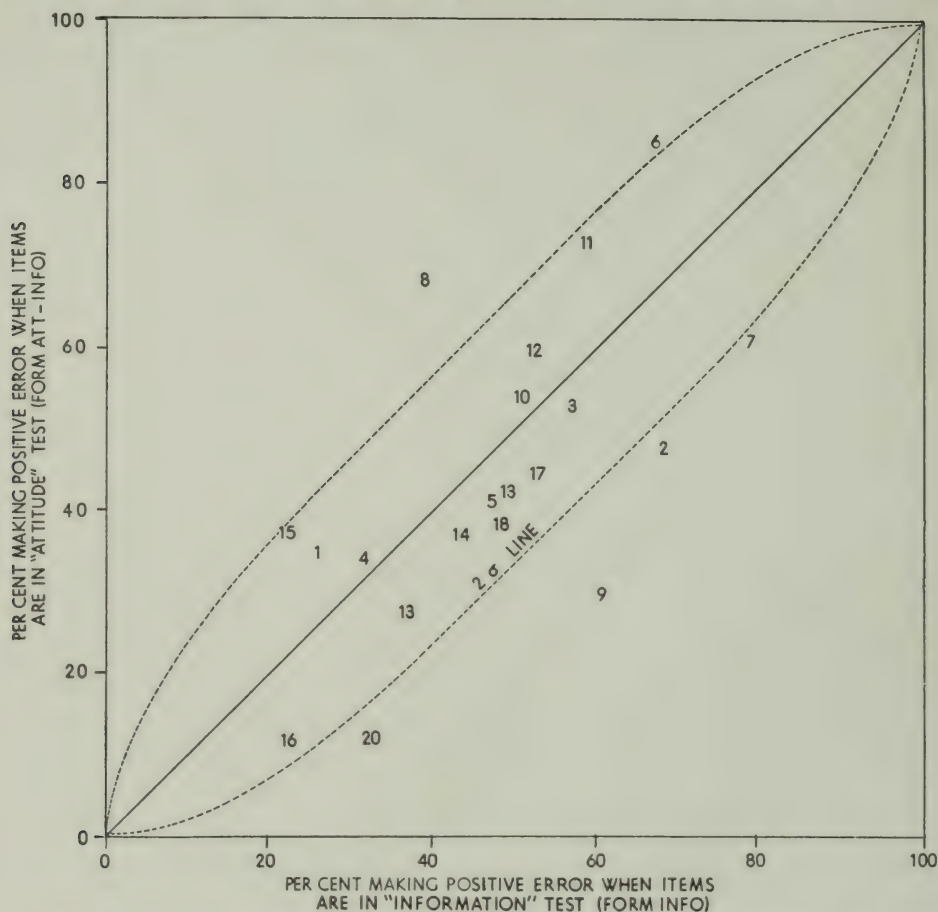


Fig. 2. Relationship between "nonfactual" items given as attitude questions and given as information questions—Labor questionnaire.

proper for a given test situation. For example, if we handed our sample a slip of paper and asked the subjects to indicate whether they were pro or anti labor, would we get the same thing a week later? Orthodox questionnaires resemble this situation in so far as they make the implication of each item clear to the subject. The error-choice technique, on the other hand, prevents control because the implications of the items are not clear. For a subject attempting control, this produces affect—and affect reduces reliability because it introduces an unstable variable.

Lest it be assumed that low reliability is a function of the items' subject matter, it should be reported that these same items presented in orthodox agree-disagree form produced reliability coefficients of .66 on the Labor questionnaire and .79 on the Russia questionnaire, corrected by the Spearman-Brown formula.

Since the error-choice form which elicited a shift in response

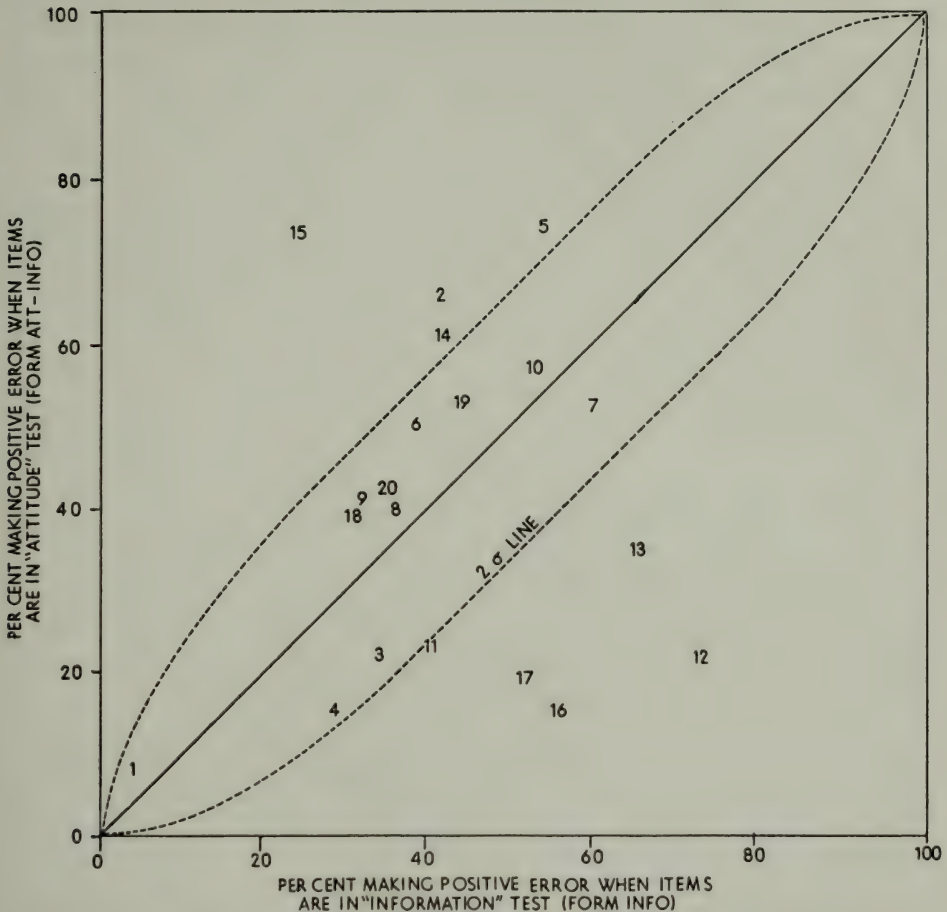


Fig. 3. Relationship between "nonfactual" items given as attitude questions and given as information questions—Russia questionnaire.

produced the higher coefficient, it is suggested that greater disguise, greater departure from orthodox methods, is the road to more meaningful high reliability coefficients.

Summary. To summarize the latter section, evidence was presented to the effect that: (1) Responses may differ for an item depending on whether that item is presented as an "attitude test" item or an "information test" item; (2) when this difference in response occurs,

variability and reliability tend to increase on the error-choice form which evokes the difference; and (3) reliability is low for error-choice forms given to a homogeneous college-student sample.

Since the error-choice technique is able to shift responses, this suggests that it is a technique adapted to the purpose of eliminating the factor of "attitude test set," inasmuch as the respondents are unaware of the implication of their error-choices. The implication of which they are unaware is that they are making the same error-choice that was demonstrated to be part of the systematic error-choice made by "known groups." The reader should not lose sight of the relationship this procedure bears to that of the projective techniques used in clinical psychology.

It is hoped that the above material offers a technique which, given a portion of the time and consideration previous methods have received, should prove useful.

16. THE RORSCHACH TEST*

Based on the interpretation of irregular blots, the Rorschach test is one of the most widely used clinical projective techniques. It is also a technique around which considerable speculation has revolved regarding its possibilities in commercial research, although little experimental work appears to have been accomplished in this area.

The present selection is useful to commercial researchers in two respects. On the one hand, it indicates the status of the test in clinical psychology and the extent of success that has been obtained with it. In addition, it advances a simple theoretical framework, supported by results of empirical tests, of the types of factors that have to be taken into account in administering the test if valid results are to be expected. The presentation is from the viewpoint of clinical psychology, but much the same considerations are applicable also to the field interviewing situation.

This article presents a theoretical frame of reference that has emerged from a number of studies of factors involved in predicting behavior from the Rorschach test. Although Rorschach devoted considerable attention to the meaning of his concepts,¹ subsequent investigators have tended in great part to ignore theory in favor of empirical studies

* Adapted from an article by Daniel R. Miller, University of Michigan, "Prediction of Behavior by Means of the Rorschach Test," *Journal of Abnormal and Social Psychology*, Vol. 48 (1953), pp. 367-75.

¹ H. Rorschach, *Psychodiagnostik* (Bern, Switzerland: Hans Huber, 1941).

of "signs" of different pathologies. As a result, there are many reports of scoring symbols that discriminate significantly between certain syndromes: depressives and normals, brain-injured and psychotics, and so forth. But as Travers points out, a test may discriminate between two groups and the theory still be wrong.² A good example is afforded by the phenomenon of "color shock" which has been shown empirically to differentiate neurotics from certain other groups. It is difficult to interpret such a finding, however, in view of the evidence that what is identified as color shock may occur in response to achromatic cards and is not always accompanied by the usual manifestations of strong emotion.³

If the principles underlying a projective test are inadequately understood, some interpretations may be valid while others should be expected to be invalid. Such is the current situation of the Rorschach test. The number of reported failures reveals that too little is known of the conditions under which the most valid results are obtained or even of what types of individuals are most accurately diagnosed. The unique level of behavior sampled by the Rorschach test can be an asset only if a theoretical system is developed to interpret responses obtained under different conditions.

This article proposes a frame of reference from which predictive variables can be identified and defined. The theory derives from a conception of the Rorschach protocol as the record of an interpersonal relationship. The dimensions of interpersonal relationships have been most thoroughly explored in therapy. Consequently, from the literature in that field were abstracted the following five categories of variables that seem pertinent to any interpersonal situation, whether it be therapy, diagnosis, or attitude measurement:

1. Setting: characteristics of the situation in which the test is taken.
2. Task: nature of the test, how it is introduced, and responses required from subject.
3. (a) Examiner's social stimulus value: readily observable characteristics such as height or certain mannerisms. (b) Examiner's character structure: ways of organizing his drives in relating to others.
4. Subject's character structure.
5. Relationship between examiner and subject: configuration of interactions.

² R. M. W. Travers, "Rational Hypothesis in the Construction of Tests," *Educational Psychology Monographs*, Vol. 41 (1950), pp. 147-212.

³ R. S. Lazarus, "The Influence of Color on the Protocol of the Rorschach Test," *Journal of Abnormal and Social Psychology*, Vol. 44 (October, 1949), pp. 506-16; J. D. Matarazzo and I. N. Menseh, "Reaction Time Characteristics of the Rorschach Test," *Journal of Consulting Psychology*, Vol. 16 (April, 1952), pp. 132-39; R. Wallen, "The Nature of Color Shock," *Journal of Abnormal and Social Psychology*, Vol. 43 (July, 1948), pp. 346-56.

It is immediately obvious that some of the five, the therapist's social stimulus value, for example, are not now formally evaluated in Rorschach interpretation. Also readily apparent is the relative interdependence of all five groupings. Task and setting are not easily separated; nor is the social stimulus value of the examiner clearly distinguished from his character structure. In order to demonstrate the derivation and possible significance of all five groupings, each will be considered in turn, first in the therapeutic relationship and then as a source of significant variance in Rorschach testing. Some exploratory investigations of interpersonal variables will then be summarized, and practical implications of the theory discussed.

Setting. Therapists generally agree that the properties of the situation must be taken into account in interpreting the patient's behavior. They have consequently devoted considerable effort to standardizing such aspects of the situation as appointment time, number of hours, number of visits per week, financial arrangements, physical position of the patient, and even the atmosphere of the room.⁴

In their interpretations of the Rorschach test, psychologists seem less concerned with the possible significance of circumstances under which the test is administered. There is, however, an increasing discussion of this topic in the literature. Schachtel feels that protocols of patients who seek help for personal difficulties may differ considerably from records of institutional inmates requested to take the test by some authority.⁵ Corroborative evidence is presented by Luchins, who interviewed a group of soldiers concerning the reasons why they gave many more responses in the testing off limits than in the standard administration. Some revealed that they had participated minimally because they felt coerced. Others said that they had taken so many tests that all had lost personal meaning for them.⁶

If a "good" record is a condition of employment or a pension, the subject may be tempted to falsify his responses. Carp and Shavzin find that some subjects can convey inaccurate pictures of their personalities under instructions to make "good" or "bad" impressions on the Rorschach test.⁷ The significance of the setting is also indicated by Kimble's

⁴ S. Freud, *Collected Papers*, Vol. II (London: Hogarth, 1946).

⁵ E. G. Schachtel, "Subjective Definitions of the Rorschach Test Situation and Their Effect on Test Performance," *Psychiatry*, Vol. 8 (November, 1945), pp. 419-48.

⁶ A. Luchins, "Situational and Attitudinal Influences on Rorschach Responses," *American Journal of Psychiatry*, Vol. 103 (May, 1947), pp. 780-84.

⁷ A. L. Carp and A. R. Shavzin, "The Susceptibility to Falsification of the Rorschach Psychodiagnostic Technique," *Journal of Consulting Psychology*, Vol. 14 (June, 1950), pp. 230-33.

finding that subjects tested under standard conditions give significantly fewer color and movement responses than others tested in the "intimate and friendly" atmosphere of a cafeteria.⁸

Task. A second major source of variables, the task, is also structured by the therapist. Macalpine is of the opinion that Freud proposed such procedures as the "basic rule," requiring that the patient report everything that comes to mind, and the therapist's silences to provide controls that would simplify the understanding of the patient's reactions.⁹ The relationship between the definition of the task and the subject's set has been carefully explored by experimental psychologists.¹⁰ Other than agreeing to the possible significance of the nature of the task,¹¹ however, clinicians have done little to apply the research findings to the administration of the Rorschach test.

Rather than stressing standardization, texts on the Rorschach technique advocate that the examiner modify the basic instructions according to the age, cultural background, and personality of the subject.¹² Apart from individual variations in method, however, there are some major differences depending upon the predilections of the clinician. Some use a trial blot,¹³ while others think it should be avoided.¹⁴ The subject may be an arm's length from a card or many feet from a large screen. In the latter instance, his protocol may be affected by such factors as visual acuity, his distance and angle from the screen, and distortions created by the lens of the projector. It is not surprising, therefore, to find certain significant differences in scoring categories obtained from the records of two matched groups, one of whom took the test individually and the other the group form.¹⁵

⁸ G. A. Kimble, "Social Influence on Rorschach Records," *Journal of Abnormal and Social Psychology*, Vol. 40 (January, 1945), pp. 89-93.

⁹ I. Macalpine, "The Development of the Transference," *Psychoanalytic Quarterly*, Vol. 19 (October, 1950), pp. 501-39.

¹⁰ J. J. Gibson, "A Critical Review of the Concept of Set in Contemporary Experimental Psychology," *Psychological Bulletin*, Vol. 38 (November, 1941), pp. 781-817.

¹¹ A. L. Benton, "The Experimental Validation of the Rorschach Test," *British Journal of Medical Psychology*, Vol. 23 (1950), pp. 45-58; W. Joel, "The Interpersonal Equation in Projective Methods," *Rorschach Research Exchange*, Vol. 13 (December, 1949), pp. 479-82; A. Wood, E. Arluck, and H. Margulies, "Report of a Group Discussion of the Rorschach Method," *Rorschach Research Exchange*, Vol. 5 (July, 1941), pp. 154-65.

¹² S. J. Beck, *Rorschach's Test. Vol. 1: Basic Processes* (New York: Grune and Stratton, 1950); B. Klopfer and D. M. Kelley, *The Rorschach Technique* (Yonkers-on-Hudson: World Book Co., 1942).

¹³ M. R. Hertz, *Frequency Tables to Be Used in Scoring Responses to the Rorschach Ink-Blot Test* (3rd ed.; Cleveland: Western Reserve University, 1946).

¹⁴ Klopfer and Kelley, *op. cit.*

¹⁵ M. Hertzman, "A Comparison of the Individual and Group Rorschach Tests," *Rorschach Research Exchange*, Vol. 6 (July, 1942), pp. 89-108.

Sometimes the subject relates to the examiner in a face-to-face situation; other times he is some distance from the examiner. Hutt finds that some disturbed psychoneurotics who give constricted records when taking the individual form give dilated records in the group where they seem to gain more security from the relative anonymity and the feeling of companionship with other patients.¹⁶

A comparison by Ogdon of group and individual Rorschach tests administered to aviation cadets shows more unrestrained expression of highly emotional material on the group form. This is attributed to the greater distance, both physical and emotional, from the examiner afforded by this test.¹⁷

Certain types of administration require that the subject report his reactions verbally; other methods entail the writing of responses. The literacy of the subject may then affect such protocols.¹⁸ A technique requiring writing may be relatively unstructured, or it may take the form of a multiple-choice test.

Practices of administering the inquiry probably vary more than any other part of the task. Some clinicians do not use the inquiry;¹⁹ some request the subject to draw his percepts;²⁰ some attempt to use roughly standardized questions;²¹ and some feel free to vary their comments depending upon their estimates of the situations.²²

Despite the extensive experimental literature linking task and set in perception, the amount of research on the possible significance of the Rorschach technique is negligible. This may be explained by a commonly held attitude that has been expressed by one clinician as follows: ". . . minor deviations in technique, in materials, in administration and scoring will matter as little (in the group form of the Rorschach test) as they do when the test is given individually."²³

Examiner's Characteristics. In addition to studying setting and task, therapists have devoted considerable attention to the relationships

¹⁶ M. L. Hutt, "Some Notes on the Usefulness of the Rorschach Method and the Rorschach as a Group Test," unpublished paper.

¹⁷ S. B. Sells, "Problems of Criteria and Validity in Diagnosis and Therapy," *Journal of Clinical Psychology*, Vol. 8 (January, 1952), pp. 23-28.

¹⁸ M. R. Harrower-Erikson and M. E. Steiner, *Large Scale Rorschach Techniques* (Springfield, Illinois: Charles C. Thomas, 1945).

¹⁹ *Ibid.*

²⁰ K. N. Levine, J. R. Grassi, and M. J. Gerson, "Hypnotically Induced Mood Changes in the Verbal and Graphic Rorschach: A Case Study," *Rorschach Research Exchange*, Vol. 7 (October, 1943), pp. 130-44.

²¹ C. Buhler, K. Buhler, and D. W. Lefever, *Rorschach Standardization Studies*, 2nd mimeographed ed., 1949.

²² Klopfer and Kelley, *op. cit.*

²³ Harrower-Erikson and Steiner, *op. cit.*, p. 8.

between their own personal characteristics and patients' behavior. The characteristics studied may be divided into two categories: the therapist's social stimulus value, and his character structure. As for the former, Oberndorf and Bibring-Lehner attribute some of their therapeutic failures to similarities in appearance, age, and social mannerisms between themselves and punitive parents of the patients.²⁴

In passing, it is interesting to note some parallel studies of the social stimulus value of the interviewer engaged in the measurement of attitudes. Robinson and Rohde find that anti-Semitic opinions tend to be withheld from interviewers of Jewish appearance or name.²⁵ A comparison has been made by Katz of the findings of white-collar interviewers of the American Institute of Public Opinion with results obtained by recruited working class interviewers. Although both staffs received the same instructions, the middle-class interviewers, apparently because of speech and dress as well as bias, elicit more conservative attitudes among lower-income groups than do the working class interviewers.²⁶

The personal problems of the therapist or interviewer may play an even more important role in his professional relationships. According to Freud, it is not possible for the therapist to replace feelings of countertransference toward the patient with neutral reactions when unresolved repressions have created "blind spots" in his perceptions. He then tends to project "as a scientific theory . . . some of the peculiarities of his own personality which he has dimly perceived."²⁷

In the field of attitude measurement, Rice studied the interviews of social investigators with selected groups of applicants for charity. Among his interviewers were the prohibitionist who reported that the difficulties in most of his cases were caused by alcohol, and the socialist who attributed the problems of his group predominantly to industrial factors.²⁸ An inquiry by Stanton and Baker into data obtained by five experienced workers indicates that the bias of the interviewer affects the outcome of the interview even when he is experienced, the direction

²⁴ C. P. Oberndorf, "Unsatisfactory Results of Psychoanalytic Therapy," *Psychoanalytic Quarterly*, Vol. 19 (July, 1950), pp. 393-407; G. Bibring-Lehner, "A Contribution to the Subject of Transference-Resistance," *International Journal of Psycho-Analysis*, Vol. 17 (April, 1936), pp. 181-89.

²⁵ D. Robinson and S. Rohde, "Two Experiments with an Anti-Semitism Poll," *Journal of Abnormal and Social Psychology*, Vol. 41 (April, 1946), pp. 136-45.

²⁶ D. Katz, "Do Interviewers Bias Poll Results," *Public Opinion Quarterly*, Vol. 6 (Summer, 1942), pp. 248-68.

²⁷ Freud, *op. cit.*, p. 330.

²⁸ S. A. Rice, "Contagious Bias in the Interview: A Methodological Note," *American Journal of Sociology*, Vol. 35 (November, 1929), pp. 420-23.

of the bias known to him, and the material has no personal or emotional connotation.²⁹

As for the Rorschach test, the few experimental investigations of examiner differences indicate that they may constitute a significant source of variance. According to the Army Air Forces, the number of responses varies with the examiner.³⁰ Lord reports that consistent differences in certain scoring categories are obtained by three examiners regardless of whether they took positive, negative, or neutral attitudes, or whether the protocol was obtained from a first or second administration of the test.³¹ Curtis and Wolf find significant differences in the numbers of overt and covert sexual responses given to female and male examiners by male veterans of World War II.³²

In none of these studies has the attempt been made to discriminate between special stimulus value and personality. Lord does attempt, however, to group the responses obtained by each examiner into a configuration from which she infers his personality. One pattern, for example, ". . . somewhat resembles what one might expect of a person confronted with a threatening, frustrating situation. If this total effect may be considered a mirror of the administrator's personality, then Examiner A would be described as a cold, forbidding, frustrating, threatening figure. . ."³³

In the absence of objective criterion measures of the examiners' personalities, Lord obtained subjective descriptions by two "psychologically sophisticated persons" who, she feels, corroborate her impressions. Examiner A is described as the "coldest, most inflexible, and most solid of the examiners."

Psychologists are becoming increasingly aware of opportunities for distortion provided by the many unstructured features of Rorschach administration and interpretation. The examiner's needs can find expression in his method of gaining rapport, phrasing of instruction, expression of encouragement, amount of pressure exerted, and his

²⁹ F. Stanton and K. H. Baker, "Interviewer Bias and the Recall of Incompletely Learned Materials," *Sociometry*, Vol. 5 (May, 1942), pp. 123-34.

³⁰ U.S. Army Air Forces Aviation Psychology Program Research Report (J. P. Guilford, ed.), *Printed Classification Tests*, Report No. 5 (Washington, D.C.: Government Printing Office, 1947).

³¹ E. Lord, "Experimentally Induced Variations in Rorschach Performance," *Psychological Monographs*, Vol. 64 (1950), No. 10, Whole No. 316.

³² H. S. Curtis and E. B. Wolf, "The Influence of the Sex of the Examiner on the Production of Sex Responses on the Rorschach," *American Psychologist*, Vol. 6 (August, 1951), pp. 345-46 (abstract).

³³ Lord, *op. cit.*, p. 27.

perceptions of the nuances of responses, the un verbalized feelings, and general behavior of the subject. The lack of structuring of the interpretive procedure allows such leeway for projection that, as Macfarlane has observed, ". . . in the hands of the clinically inexperienced, the doctinaire, or the methodologically uniformed [the Rorschach test] easily degenerates into nothing but one more projective tool—to wit, one that discloses the organizing dynamics of the interpreter rather than the organizing dynamics of the research subject."³⁴

It seems probable that the interpretations of even sophisticated clinicians express their personal dynamics to some degree. Frank attributes the great difficulty he has had in comparing the reports of a number of competent clinicians to the fact that each interpreter warps the material in accordance with his or her own clinical training and personality.³⁵ In a study of the assessment of clinical skills, Kelly and Fiske find that the sum of movement responses correlates .20 with the criterion of clinical proficiency, while a rating based on the global analysis of the Rorschach test correlates .04. They attribute the superior predictive value of some of the determinants to the fact that their greater objectivity permits less projection than the global measure.³⁶

Subject's Reactions. The subject's behavior provides a fourth grouping of variables. Of primary concern to the therapist and test administrator are those aspects of the subject's behavior that are representative of his everyday relationships with other significant people. These generalized behavioral tendencies are interpreted in terms of such indices as different types and degrees of control, degree of anxiety, intellectual level, and the expression of emotion. Care should be taken to isolate temporary states. Different moods suggested during hypnosis are accompanied by corresponding changes in the protocols.³⁷ Apparently some of the cards are more sensitive to the emotional changes that accompany the administration of a drug than others.³⁸ Subjects described as overly cautious or fearful in the normal waking

³⁴ J. W. Macfarlane, "Problems of Validation Inherent in Projective Methods," *American Journal of Orthopsychiatry*, Vol. 12 (July, 1942), p. 405.

³⁵ L. K. Frank, "Understanding the Individual through Projective Techniques," *American Council on Education Studies*, Vol. 14, Series I, pp. 56-62.

³⁶ E. L. Kelly and D. W. Fiske, *The Prediction of Performance in Clinical Psychology* (Ann Arbor: University of Michigan Press, 1951).

³⁷ Levine, Grassi, and Gerson, *op. cit.*; T. R. Sarbin, "Rorschach Patterns under Hypnosis," *American Journal of Orthopsychiatry*, Vol. 9 (April, 1939), pp. 315-18.

³⁸ C. N. Cofer, "Psychological Test Performance under Hyoscine: A Case of Post-Infectious Encephalopathy," *Journal of General Psychology*, Vol. 36 (April, 1947), pp. 221-28.

state give more productive Rorschach records when hypnotized or under sodium amytal.³⁹ Temporary experimental inhibition of motor activity is followed by an increase in movement responses on the Rorschach test.⁴⁰ It is possible that temporary frustrations, inhibitions, and gratifications may affect the subject's pattern of responses in ways that are not typical of his usual behavior.

Relationship. This fifth group of variables defines interaction between the participants. Jung represents almost all schools of therapy in his statement that the therapist does not "analyze" an object at a distance, but is as much a part of the relationship as the patient. Both are mutually influenced in the treatment.⁴¹ According to Rogers, ". . . to treat another person as a person is to open oneself to change through the influence of the relationship."⁴² Sullivan regards the therapist as a constituent element in an interpersonal situation, a "participant observer."⁴³

To date, this theory has not been explored experimentally in connection with projective techniques. Its application to the Rorschach test introduces the possibility of a continually changing series of interactions between examiner and subject. The nature and significance of these interactions may be illustrated by a hypothetical relationship between a dominant clinician who becomes anxious unless there is psychological distance between himself and others, and a dependent patient who becomes disturbed when he lacks support.

In the presence of the assertive examiner, the passive individual may respond warmly and produce a dilated record on his first few cards. If this expression of warmth makes the examiner anxious, he may defend himself by becoming aloof. The retreat may elicit the subject's conscious resentment and his Rorschach responses may begin to show signs of hostility. After a while he may repress his antagonism, and produce a constricted neutral record on the last few cards. Finally, the examiner may express his anxiety during the inquiry by alternating between too much and too little questioning.

³⁹ W. L. Wilkins and A. J. Adams, "The Use of the Rorschach Test under Hypnosis and under Sodium Amytal in Military Psychiatry," *Journal of General Psychology*, Vol. 36 (April, 1947), pp. 131-38.

⁴⁰ S. J. Korchin, J. Meltzoff, and J. L. Singer, "Motor Inhibition and Rorschach Movement Responses," *American Psychologist*, Vol. 6 (August, 1951), pp. 344-45 (abstract).

⁴¹ J. Jacobi, *The Psychology of C. G. Jung* (New Haven: Yale University Press, 1951).

⁴² C. R. Rogers, "Where Are We Going in Psychology?" *Journal of Consulting Psychology*, Vol. 15 (June, 1951), pp. 171-77.

⁴³ H. S. Sullivan, "Conceptions of Modern Psychiatry," *Psychiatry*, Vol. 3 (February, 1940), pp. 1-117.

If these speculations are correct, the continuously changing interpersonal process reflected in the Rorschach protocol is comparable in complexity to the relationship that occurs in a therapeutic interview. Interpretations are therefore grossly oversimplified when phrased in terms of static classificatory concepts that ignore shifts in the relationship. In addition to such categories as *Erlebnistyp* and pathognomonic responses, the clinician's report should also consider the sequence of responses in a frame of reference consisting of all five groupings of variables. The complex interactions that might occur among the five areas are indicated in Figure 1, which is a modification of Sullivan's diagram of the interpersonal situation.⁴⁴

The subject responds verbally to the task (T) in a structured situation (S) in the presence of an examiner with certain characteristics

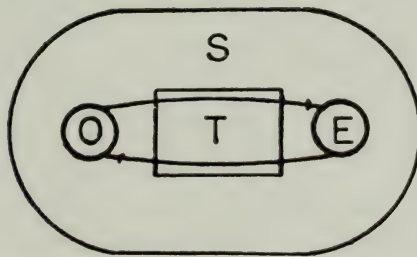


Fig. 1. Interpersonal situation during administration of the Rorschach test.

(E) to whom the subject relates (arrows) in terms of his organization of drives in interpersonal situations (O). An understanding of the protocol necessitates an examination not only of these variables, but also of the perceptions and emotional reactions that they arouse in examiner and subject.

The protocol may also reflect certain interpersonal factors that are not directly dependent upon the test. One clinician, for example, may regard the setting as a reflection of his professional status; to his subject the setting may signify a hurdle to the attainment of a job. Another examiner may be oriented to the solution of an interesting problem, while his subject may anticipate finding out why he is so nervous. The task may intrigue one examiner because it yields such rich materials; his patient may feel threatened because of its relative lack of structure. Another subject may use the test as a cathartic instrument in the presence of an examiner who is bored because he has been working too many hours. The subject's overtures may seem

⁴⁴ H. S. Sullivan, "The Meaning of Anxiety in Psychiatry and in Life," *Psychiatry*, Vol. 11 (February, 1948), pp. 1-13.

friendly to one examiner who attempts to give support. Another feels the subject too "forward" and becomes reserved. In response to the examiner's support in the inquiry, one individual may relax and give many new responses. Another may become embarrassed and give a constricted record. The subject's record may contain covert signs of hostility because the examiner's age and appearance are similar to those of his authoritarian father. If these same qualities are reminiscent of a previous teacher, the subject may give a productive record as he would in an intelligence test.

EMPIRICAL INVESTIGATIONS

At the university of Michigan, a number of studies have been devoted to preliminary explorations of some variables in the five groups. Sanders and Cleveland studied the relationships between examiners' personalities and the Rorschach records of their subjects. The examiners administered the test under highly standardized conditions to randomly selected male sophomores. After taking the test the subjects filled out check-lists concerning the degrees of overt hostility and anxiety that they sensed in the clinicians.⁴⁵ To secure covert measures of the examiners' hostility and anxiety, their personal Rorschach tests were scored by means of Elizur's technique of content analysis.⁴⁶ Significant over-all differences among examiners are reported for 20 of 38 scoring categories. The epsilons indicate that these differences account for 3 to 7 per cent of the variances of significant scores. On each of the overt and covert measures of hostility and anxiety, examiners in the upper and lower thirds obtain significantly different protocols.

Rorschach records of subjects given different test definitions and ego-involving instructions were compared by Calden and Cohen. A questionnaire filled out by the subjects was used to check degree and type of ego-involvement. Significant over-all variances ranging from 3 to 10 per cent were found among the different experimental groups for 17 of 25 conventional scores and 7 of 8 percentage scores. In short, protocols tended to vary depending upon whether the test was defined to the subject in terms of imagination, nervousness, or intelli-

⁴⁵ R. Sanders, *The Relationship between Examiner Hostility and Subjects' Rorschach Scores* (Ph.D. dissertation, University of Michigan, 1950); S. E. Cleveland, *The Relationship between Examiner Anxiety and Subjects' Rorschach Scores* (Ph.D. dissertation, University of Michigan, 1950).

⁴⁶ A. Elizur, "Content Analysis of the Rorschach with Regard to Anxiety and Hostility," *Rorschach Research Exchange*, Vol. 13 (September, 1949), pp. 247-84.

gence, whether the ego-involvement was high or low, and whether it was externally or self-imposed.⁴⁷

Filer investigated the relationships between clinicians' personalities and their interpretations of projective tests. He tabulated the relative frequencies of references to aggression, hostility turned inward, passivity-dependency, and feelings of insecurity in the reports of 13 examiners. These diagnostic categories were combined into objective patterns. Relationships were then postulated between diagnostic patterns in the examiners' reports and ratings of their actual behavior. Successful predictions were made of examiners' positions on the dimensions of ascendancy, depression, and extrapunitive-ness, intropunitive-ness, and impunitiveness.⁴⁸

In a second part of the research, Filer tabulated the average frequencies of references made by the entire group of examiners to 13 different defenses. A group of judges then ranked each defense according to the degree that it would be operative in the personalities of clinicians and of patients. These rankings were then compared to the frequencies of references to defenses in the examiners' reports.

The three defenses mentioned most frequently are intellectualization, compulsivity, and conscious control. All three are ranked by the judges in the lower half of the patients' defensive hierarchies, but in the upper half of the psychologists' defensive hierarchies. In other words, the three defenses that appear most often in the reports are more characteristic of the examiners' personalities than of their subjects. However, there is more than mere projection in the interpretations. Of the defenses ranking fourth to ninth in the psychologists' reports, five coincide with the first five in the judges' rankings of the patients.

Sources of Error. These experiments point to a need for investigations of all the sources of variance on the Rorschach test so that the conditions under which it is most and least valid may be defined. The highest validities of standard intelligence tests, for example, have been found with subjects who are white,⁴⁹ middle-class,⁵⁰ without

⁴⁷ G. Calden, *The Relationship of Varied Test Definitions and Degrees of Ego-Involvement to Rorschach Test Performance* (Ph.D. dissertation, University of Michigan, 1951); L. Cohen, *The Role of Two Attitudinal Factors on Group Rorschach Performance* (Ph.D. dissertation, University of Michigan, 1951).

⁴⁸ R. N. Filer, *The Clinician's Personality and His Case Reports* (Ph.D. dissertation, University of Michigan, 1951).

⁴⁹ B. F. Haught, "Mental Growth of the Southwestern Indian," *Journal of Applied Psychology*, Vol. 18 (February, 1934), pp. 137-42; O. Klineberg, *Negro Intelligence and Selective Migration* (New York: Columbia University Press, 1935).

⁵⁰ K. Eells, et al., *Intelligence and Cultural Differences* (Chicago: University of Chicago Press, 1951).

personality disorder,⁵¹ Northern,⁵² and urban.⁵³ Of particular importance to research on the Rorschach test⁵⁴ are factors, such as those in the following list, that seem related to the subject's responses but are not typical of his ways of relating to people.

A. Setting:

1. Interaction between receptionist and subject.
2. Physical characteristics of examining room such as color, lighting, anatomical charts, or medical instruments.
3. Subject's reaction to examiner's use of stop watch and detailed note-taking.
4. Atmosphere—experimental laboratory, psychiatric ward, or employment office.

B. Task:

1. Reasons given subject for taking test.
2. Wording of instructions.
3. Ego-involvement in test.
4. Background factors, such as social class, that might predispose subjects toward certain sets.

C. Examiner:

1. Appearance.
2. Attitudes.
3. Needs.
4. Defenses.

D. Subject:

1. Current mood.
2. Temporary physiological state.
3. Perceptions of unique aspects of the testing situation.
4. Special disabilities.

E. Relationships—interactions between examiners of different character structures and various types of subjects:

1. The overambitious clinician who needs "rich" records and the constricted depressed patient.
2. The cold, objective scientist and the intellectual student.
3. The authoritarian examiner and the phobic character.
4. The overly solicitous clinician who evaluates himself in terms of his success in obtaining rapport and the case of traumatic neurosis.

⁵¹ C. H. Altman and D. Shakow, "A Comparison of the Performance of Matched Groups of Schizophrenic Patients, Normal Subjects, and Delinquent Subjects on Some Aspects of the Stanford-Binet," *Journal of Educational Psychology*, Vol. 28 (October, 1937), pp. 519-29.

⁵² Klineberg, *op. cit.*

⁵³ H. E. Jones, H. S. Conrad and M. B. Blanchard, "Environmental Handicap in Mental Test Performance," *University of California Publications in Psychology*, Vol. 5 (January, 1932), pp. 63-99.

⁵⁴ Studies are also needed of comparable sources of error on objective techniques, such as the Stanford-Binet test and the Minnesota Multiphasic Personality Inventory, and less structured projective instruments, such as the Draw-a-Man and Thematic Apperception Tests.

SUGGESTIONS TO THE PRACTICING CLINICIAN

What is the practicing clinician to do in the absence of this information? In general, it seems advisable for him to plan his testing so as to control the different sources of error. Controls can often be inserted into the setting. Care can be taken, for example, to insure that all referring therapists in a clinic give their patients similar explanations for the testing. A more difficult problem is provided by the variables that cannot be controlled in this way. The natures and magnitudes of such variables must be ascertained so they can be used as corrective factors in the final interpretation of obtained results. Possible methods of aiding the clinician to take sources of error into account tend to fall into three areas: the subject's perceptions, the adequacy of the behavioral sample, and the examiner's distortions.

The Subject's Perceptions. After administering the Rorschach test, Schachtel questions the subject in detail concerning his conception of the nature of the test and his attitudes toward the examiner. The results of the interview are then used both for purposes of control and as additional samples of behavior. Calden and Cohen's studies suggest that some of the information sought by Schachtel could be obtained by an objective questionnaire.⁵⁵

Adequacy of Behavioral Sample. A second issue that should concern the clinician is the representativeness of the subject's responses. It is important to acquire, in addition to the subject's point of view, a representative sample of his interpersonal responses. Current clinical practice often entails the prediction of a patient's future behavior in many kinds of therapeutic situations from his reactions to one test administered by one examiner in one setting. It is likely that estimates of diagnosis and prognosis would improve considerably if a number of tests were administered by different examiners at different times. The more varied the levels measured by the techniques, the greater would be the probability that the subject's reactions in the clinical setting were representative of his behavior in other kinds of situations. If results of different tests agreed or were congruent, the clinician could be more certain of his interpretation than when he used only one test. If the data were incompatible, however, he could not know which tests, if any, were valid. From one point of view, this might be an advantage. It seems wiser for the diagnostician to report that

⁵⁵ Calden, *op. cit.*; Cohen, *op. cit.*

the data are equivocal, than for him to risk misleading the therapist by incautious interpretations of nonrepresentative or fragmentary results.

Another approach to the problem of obtaining a sample of behavior is the exploration by the examiner of the limits of obtained reactions. Among the more promising techniques are the testing of limits,⁵⁶ the analysis of the subject's free associations to his responses,⁵⁷ and the measurement of his ability to alter his responses to the test when the scoring symbols are described to him.⁵⁸

The Examiner's Distortions. The validity of the Rorschach test will probably increase as the clinician gains insight into the narrowness of some of his own perspectives. An examination of the interpretations of any group of clinicians, for example, indicates that each tends to discuss certain topics and omit others in most of his reports. Such oversights might be partially corrected by the use of a standard inventory of topics to be covered in reporting pertinent results. This should help to prevent the kind of distortion that results from accidental oversight or selective inattention.

The susceptibility of the interpretative process to the examiner's projections also suggests that he seek the independent viewpoints of other competent clinicians before writing reports. A more objective method is suggested by the correction of individual bias in rating scales. Just as the rater's distortion is measured by the calculation of his central tendency and range, so the clinician might tabulate the frequencies of his obtained scoring categories, as well as his references to different personality variables and diagnostic categories in reports. By comparing his distributions with those of other competent examiners, the clinician could become aware of unique properties of his social stimulus value and some of his defensive distortions. This knowledge could then help to prevent some of the more common inaccuracies that stem from the unique characteristics of the examiner.

In the long run, the prediction of behavior from projective tests requires more than the palliatives described above. What is essential is a deductive system relating the findings of different investigators. The proposed theoretical frame of reference suggests many related

⁵⁶ M. L. Hutt and J. Shor, "Rationale for Routine Rorschach 'Testing-the-Limits,'" *Rorschach Research Exchange*, Vol. 10 (June, 1946), pp. 70-76.

⁵⁷ G. R. Kamman, "The Rorschach Method as a Therapeutic Agent," *American Journal of Orthopsychiatry*, Vol. 14 (January, 1944), pp. 21-27.

⁵⁸ M. L. Hutt, R. Gibby, E. O. Milton, and K. Pottharst, "The Effect of Varied Experimental 'Sets' upon Rorschach Test Performance," *Journal of Projective Techniques*, Vol. 14 (June, 1950), pp. 181-87.

testable hypotheses that might help in establishing the groundwork of such a system.

17. CARTOON AND PICTURE DEVICES*

Here are some successful examples of the manner in which the Picture Frustration and the Thematic Apperception Test have been used to solve marketing problems. Note the emphasis placed on proper sampling techniques even though psychological methods constitute the basic approach.

THE PROBLEM

One of the difficulties in marketing research is getting from the consumer an adequate answer to "why" questions. A typical procedure is to use an open-end question such as: "Why do you prefer X brand of product?" Answers will usually relate to price, quality, convenience, and other general attributes. The interviewer may follow the question up with requests for further information by using phrases such as: "Tell me more," or "What do you mean?" A great deal of the responsibility for the success of this type of interview depends upon the background of the interviewer as well as his skill. Much time in general conversation may be needed to bring the respondent to the subject at hand. Picture Frustration and the Thematic Apperception Test are two projective techniques that may be used for obtaining answers to "why" questions. It is the purpose of this paper to discuss their application to such problems.

Picture Frustration (PF). PF, as devised by Rosenzweig, consists of 24 pictures of the cartoon variety.¹ Each picture represents a different situation; there is no continuity of ideas among the pictures. Two figures are included in each cartoon. Facial features and other individualizing characteristics are omitted in order to render the figures more neutral. One of the figures says something which is potentially frustrating to the other. The speech balloon for the frustrated person is left blank. The respondent is instructed to fill this in, using the very

* Adapted from an article by Martin Zober, Drake University, "Some Projective Techniques Applied to Marketing Research," *Journal of Marketing*, Vol. 20 (January, 1956), pp. 262-68.

¹ S. Rosenzweig, "The Picture-Association Method and Its Application in a Study of Reactions to Frustration," *Journal of Personality*, Vol. 14 (September, 1945), pp. 3-23.

first words that come to his mind. Presumably he will project his own biases into the situation.

Thematic Apperception Test (TAT). TAT, developed by Henry A. Murray, consists of 20 cards on which are printed pictures carefully selected from magazine illustrations, paintings, drawings, and other art sources.² The respondent is asked to make up a story around each picture, telling what events led up to the situation, what the outcome will be, and describing the thoughts, feelings, or personalities of the characters.³

PF AND TAT IN MARKETING RESEARCH

Although the original PF and TAT situations illustrated were potentially frustrating to the respondent, pictures in marketing research do not necessarily have to be designed about a frustrating situation. The respondent may have to resolve a problem or make a decision. The important aspect of the PF technique is to get the respondent to project himself into the situation, revealing his attitude toward the situation by talking about the character in the picture. Examples in this article illustrate the use of the technique.

Designing PF and TAT Situations. Before a PF or TAT situation can be designed, a series of depth interviews should be made in the field to get the feel of the problem. It is then possible to develop problem areas about which pictures may be created. In designing picture situations, there should be consultation with a person trained in social research. An artist may then be used to draw the situation, using cartoon style. The characters should be done without definite expressions and should not be ludicrous, since the respondent must project himself into the situation. The use of multilith stencils permits the artist to use a free form and to make corrections easily.

After the picture is drawn, it must be tested in the field. Several interviews should be made to determine whether the responses elicited are related to the subject being studied. For example, two of the pictures which were originally designed for the Chamber of Commerce survey discussed below had to be revised after testing was done. One picture situation that showed a man and woman at a restaurant eating was discarded because the picture of the man holding a glass of wine while talking proved to be socially offensive to the respondents.

² H. A. Murray, et al., *Explorations in Personality* (New York: Oxford University Press, 1938).

³ Examples of TAT and PF are given by G. H. Smith, *Motivation Research in Advertising and Marketing* (New York: McGraw-Hill Book Co., Inc., 1954).

Another picture showed a woman talking on the telephone. She said, "I am going to X City tomorrow . . ." The responses covered such a wide range of subjects that the picture was found to be unsuitable.

In using the pictures, the interviewer should hand the pictures one at a time to the respondent. The interviewer should read the statement in the balloon and then ask, "What does the other woman say?" In general, the respondents may find the first picture strange and the interviewer may have to repeat the procedure and reassure the respondent that there is no "correct" answer. Once a response is obtained to the first picture, there is usually no problem in getting responses to the other pictures. It is most important that the interviewer suggest no possibilities to the respondent; the responses must be the things which come to the respondent's mind.

The revised pictures should be tested again and corrected until they are satisfactory in producing the type of information that relates to the situation being studied.

Sampling Technique. In all surveys where TAT or PF pictures were used, either a stratified or random sample was used. In the Chamber of Commerce survey, the Politz not-at-home technique was used to save time on call backs. The answers to projective questions are subject to the normal sampling errors, assuming that random selection is in fact achieved. Usually, only a few individuals do not respond to the PF or TAT technique. In the Chamber of Commerce survey, only two of the four hundred persons interviewed failed to respond to the picture technique. A reduction of sample size of this magnitude is insignificant. All the rules of good sampling procedure should be followed in the field work. No interviewer should have the opportunity to select the house or person he will interview.

Tabulation. The simplest way to tabulate responses from a picture interview is to have all the responses typed on cards appropriate to the dimensions of the responses. This facilitates reading the material and makes it easier to group it into logical categories. In grouping the responses, it is important to classify references made to any subject area pertinent to the survey. For example, a response might be: "I like to get good bargains downtown, but I do dislike the crowds." Two specific references are made: one to the fact that this person likes bargains (there might be another group that dislikes bargains); the second reference is to the fact that he dislikes crowds (there might be another group that likes crowds). In general, a definite pattern of responses develops.

Analysis Based on Pictures. Although the pictures may lead to certain hypotheses, the hypotheses must be related to the results of other questions in the questionnaire. The pictures will give a finer qualitative interpretation to the other questions. Pictures may be used with such other devices as sentence completion, open-end questions, and word association.

CASE HISTORIES

A Grocery Store. A grocery store was experiencing a decline in sales. The front of the store was modernized, a cooler for vegetables was installed, and the premises were made neat and clean. The sales decline persisted, however. On the basis of discussions with management, depth interviews, and observations made in the store, the areas to be explored were decided upon.

One of the areas was the importance of prices to the consumer in his selection of a grocery store. The picture that was ultimately designed showed a woman sitting at a table drinking coffee with her neighbor. One woman said: "Well, I feel I have to buy food where the price is lower—that's the main thing as far as I'm concerned." The other woman answered: "Art and I agree that I should shop where . . ." The respondent was asked to supply the rest of the statement.

The second cartoon was concerned with the purchase of fruits and vegetables; the third, with the influence on the consumer of a remodeled store. The fourth was concerned with buying motives for meat. The fifth was concerned with the attitude toward shopping in more than one store for food. The seventh picture used the TAT technique. The respondent was shown a sketch of an interior of a grocery store and was asked to tell a story about the picture.

As an example of method of tabulation, the picture concerned with price may be used. Most of the responses to this picture made references to the fact that they preferred both a good price and quality merchandise. Here are two typical remarks illustrating this category:

Best quality for your money . . . although I don't think a bargain is a bargain lots of times . . . if I have to throw it out it's no bargain . . . it must be edible.

. . . get it by brand and where quality and price are reasonable . . . you pay a price for quality.

The second largest group of responses related to quality of merchandise only. They said:

I shop where I get better quality . . . I'd rather pay a bit more and get good stuff.

Well, I buy even if it is a little higher if we like it. We have our kinds that we like best.

Slightly below this group was a group which indicated that price was important. Here are typical responses made in this group:

. . . where the price is lowest. I believe United prices are lowest . . . big concern and can afford to have lower prices.

. . . where the food is cheapest.

Other subjects mentioned related to the following areas: similarity of prices in all stores, cleanliness of the store, convenience of parking, variety of food sold, local loyalty, and reciprocity among tradesmen.

A tabulation was run on each subject mentioned as well as the number of times a specific store was mentioned. The tabulation showed that the store which had the highest number of sales was mentioned positively most frequently.

The analysis of all the pictures in combination with the other questions produced a definite pattern. It was discovered that the most important reason for the selection of a grocery store was the reputation it had for the quality of meat sold. The store having the leading sales also led in the impression that it carried the highest quality meat. The store under investigation had a bad reputation in regard to the meat it sold. The following typical remarks were made about it:

Meat at X needs to be improved considerably. X has a bad name because of their (sic) meat department.

. . . shopped at X because of their meat . . . then stopped because meat went down in quality.

Investigation into company policy when the store was opened showed that the store did sell meat of a variable quality. Thus, the first impression on the customers in the area was a bad one. The customers who were lost did not know that the store sold high-quality meat. As a result of the survey, the store management decided that the consumers must be informed of the fact that they carried high-quality meat. This meant promotion in the newspaper and in the store windows.

Another discovery made was that the people living in the area objected to the store's advertising because it looked like the ads used by the chain's city stores. The ads were individualized to meet this criticism. The arrangement of stock was changed because of the survey's findings that the respondents objected to crowded aisles which they associated with the store. After the changes were made, the store showed an increase in sales.

Milk Company. A milk company president felt that the people of the city had a very unfavorable attitude toward his company. His information came to him from his sales force who reported that they could not compete successfully because their milk was thought to be low in butterfat and unclean. The president was planning a major change in company policy. He was not certain that the unfavorable attitude existed and how it might influence the new policy. A survey was taken to determine the attitude of the people toward the milk company.

After depth interviewing and testing picture situations, the following picture situations were used:

1. Scene: Two women sitting in a living room. First woman: "Our Junior insists we get the milk that is advertised to make him strong." Second woman: (Balloon was left blank).

2. Scene: Two women talking over the backyard fence. First woman: "Mrs. Jones just changed milk companies. Are you going to change too?" Second woman: (Balloon is left blank).

3. Two women with baby carriages meet on the street. First woman: "We buy the milk that's cheapest because every penny counts when you have a family." Second woman: (Balloon is left blank).

4. Scene: Two women in a grocery store. One is pushing a cart. The other is bending over a large refrigerated case. First woman: "Would you get me a bottle of milk while you're bending down—any brand—they're all the same." Second woman: (Balloon is left blank).

5. Scene: Two women in a kitchen. First woman: "My, this milk tastes good! Which company do you buy from?" Second woman: ". . . company because . . ."

6. Scene: A milkman delivering a bottle of milk to a housewife. Milkman: "Welcome to our neighborhood! I hope you'll be a customer for our milk because . . ."

As an example of the analysis, the responses to the first picture may be used. The picture tested the importance of children's choice in the selection of a brand of milk used. It was found that the mothers were equally divided in the matter. Those who were influenced by their children said:

. . . Junior insists we get milk from that nice man who wears the cowboy hat with his uniform.

. . . My boy does too. He saw an ad of Hopalong Cassidy and decided he wanted to be big and strong like him . . .

My son insists on A. E. because he belongs to the Kiddie Klub on the radio every morning . . .

Most of the responses to the picture related to the quality of milk and the fact that there was no difference in quality among various milk companies. To use the words of a respondent:

All milk is processed under the same state laws so one isn't much better than the other . . .

The main conclusion of this survey was that the milk company did not have an unfavorable reputation in the minds of the respondents. On the basis of this finding, the company launched a new pricing policy and has been very successful with it.

Chamber of Commerce. The Chamber of Commerce was interested in finding out how the merchants of the city might improve their services to the consumers coming in from outside of the city.

The following picture situations were developed after depth interviews and tests:

1. Scene: Two women sitting in the kitchen drinking coffee. First woman: "Ed's organization is having a convention in X City and I can't make up my mind whether or not to go along." Second woman: "When Henry's organization meets, I . . ."

2. Scene: Two women are shown on a street with baby carriages in front of them. First woman: "The last time I shopped in X City the clerk who waited on me . . ."

3. Scene: Large room suggesting church basement where two women were setting a long table. First woman: "Mary, you'll have to come over and see the bargains I got in X City!" Second woman: "When I go to X City . . ."

4. Scene: Two women standing in front of the Art Center. First woman: "This is my first visit to the Art Center. I'm planning to take up sketching here." Second woman: "I'm glad you can do it. For me the cultural opportunities in X City . . ."

5. Scene: A man and a woman standing in front of a furniture store. Man: "Grace, now that we're in X City, what about stopping in here?" Woman: (Balloon left blank).

6. Scene: Two women standing in a grocery store in front of their grocery carts. First woman: "Mrs. Jones told me she's not going to buy at our local dress shop any more because she knows the clerks too well." Second woman: (Balloon left blank).

7. Scene: Two women talking over the back fence. First woman: "You know that stuff I bought in X City. I don't think it's exactly what I want." Second woman: (Balloon left blank).

8. Scene: A man and woman seated in front of the television set in their home. Woman: "I see there are some big shows coming to X City, honey—Pajama Game, Martin and Lewis, the Sports Motor Cavalcade. What do you say—shall we go to X City?" Man: (Balloon left blank).

9. Scene: Women seated around a card table in a home. First woman: "I went to that store in X City. The right people shop there." Second woman: "I like to shop . . ."

10. Scene: Two women standing in front of a post office which looks like a small-town post office. First woman: "I was looking for a gift for Ethel and Harry, but I can't seem to find the right thing." Second woman: "The shops in X City . . ."

11. Scene: A woman is shown in a car near a sign along the road which says "X City—10 miles." Interviewer: "Tell me a story about this picture."

12. Scene: A department store interior showing a crowd around a table with the word "sale" above it. Interviewer: "Tell me a story about this picture."

An analysis of the picture for Scene 11, showing a woman in a car near a sign along the road which said "X City—10 miles," used the TAT technique. Most of the respondents had an immediate reaction that they would be happy to be that close to their destination.

Here are some of their comments:

She is happy the ride's over . . . relieved to get there . . . tired.

Oh boy, only ten more miles . . . I can hardly wait.

I'd be tickled if I was just ten miles from X City. I'd soon be there.

I sure am glad to get there . . . hope to get a good place to park. Turn me loose there; I won't want to go anywhere else.

From these remarks, it is apparent that most of the respondents traveled a distance greater than ten miles and that the drive is a chore.

Some of the people responded with remarks showing apprehension about city traffic. They said:

This is where I start getting lost . . .

I'm more conscious of the traffic. Roads are very bad in X City . . .

Some of the answers related to things that people come to X City to do. The one mentioned most frequently was shopping. Statements such as the following were common:

The girl is evidently shopping-bent and happy to be near X City.

Other remarks related to the food they would have and where they would meet or eat. For example:

I think of where we are going to eat when we get there.

She says let's all meet at Younkers for lunch . . . it's been a long way.

. . . I want to go to the Utica but we'll all meet at the bus depot at four o'clock.

SUMMARY AND CONCLUSION

The theory of PF and TAT is to get the respondent to project himself into a situation that calls for some type of decision. In PF the respondent is led into the situation by copy in a balloon over one of the characters in the picture. In TAT the respondent is asked to tell a story about the picture. In both cases the interviewee projects himself into the situation and tells how he feels about it.

Although the picture responses may lead to conclusions, it is important to relate the conclusions to the results of the objective questions in the questionnaire. The pictures may give a finer qualitative interpretation of the objective questions or they may explain material found in other parts of the questionnaire where such techniques as sentence completion or word association may be used. As noted, standard marketing research practices must be used with PF and TAT pictures. The pictures should be designed after a series of depth interviews have been held.

As few as three pictures and as many as twelve may be used. When there is a lengthy questionnaire, the number of pictures should be reduced. The picture changes the tempo of the interview and restimulates responses to a subject in a different way. Questionnaires using pictures have lasted as long as an hour and as little as twenty minutes.

The sample size in the examples given above varied from 94 to 400 interviews. The size of the sample depends upon the universe to be surveyed, the degree of stratification, and the maximum error that could be allowed. Generally, when dealing with qualitative attributes, allowance must be made for the type of error implicit in the method. There is often, however, a strong concentration of responses in a particular subject area. In general, the responses tend to follow a pattern where extremes are expressed by a few, and most respondents express attitudes between the extreme poles.

Instructions must be given to the interviewer regarding interviewing procedure. Interviewers may then do sample interviews among each other during the training session.

Ehle cites the following advantages of the cartoons:⁴

1. Interviewees seemed to enjoy working with pictures; rapport was facilitated.
2. Pictures were found to break the ice for difficult or irritating items.
3. The pictures may substitute for written questions in cases where respondents find it difficult to talk about their attitudes toward figures, their personal involvements, self-derogatory attitudes, and the like.
4. Finally, pictures of the projective variety may yield more information than can be gotten by direct questioning.

Neither TAT nor PF eliminates the responsibility of the marketing research analyst to find the relationships of the data he has and to draw valid inferences. All the biases that might occur in the sample, the questionnaire, the interviewer, and the analyst are present—and

⁴ E. L. Ehle, "Techniques for Study of Leadership," *Public Opinion Quarterly*, Vol. 13 (Summer, 1949), pp. 235-40.

perhaps to a greater degree than in straight quantitative analysis. The use of pictures is valid, however, and often enhances the effectiveness of other techniques in getting consumer motivation.

18. THE MEANING OF GASOLINE SYMBOLS*

Here is another highly illuminating example of how motivation research is carried out, this study relying on a different battery of techniques than the preceding one. Some of the tests used in this study are of the projective variety, while others are not. The study as a whole shows how different tests can serve to supplement each other and remedy, to some extent, their separate deficiencies.

It is a general practice of gasoline companies, as well as of many other companies, to utilize some general sign and emblem as a part of the trademark or brand identification. These symbols, as they will hereafter be called, are used in advertising, in gas station signs, and in other ways to identify the product and brand offered for sale.

Companies differ in the extent to which these symbols are used and in the extent to which they are made a central, as opposed to a casual, part of the companies' public notices. It is generally assumed that such signs will connote for the public the general notion of "gasoline" as well as more specifically denote the particular brand in question. It is, of course, generally true that these symbols seldom occur independently of further indication of the particular brand. Thus, the complete symbols in each case might be thought to be the particular emblem plus the brand name of the gasoline. These symbols tend to be larger in size and more noticeable, however, than the written brand name which generally accompanies them. The company thus hopes that the symbol will in itself draw the attention of the public. It further hopes that the symbol will be a positive and encouraging factor in drawing the attention of the public—as opposed to a negative and discouraging factor and as further opposed to a merely neutral factor.

It is thus possible to consider the symbol as being in some fashion specifically denotative of a particular brand of gasoline, and generally

* This is part of a more general study on gasoline marketing carried out by William Henry, Chairman, Committee on Human Development, University of Chicago, for the *Chicago Tribune* and reproduced with the kind permission of both author and sponsor.

denotative of gasoline of no particular brand, or as merely an aesthetic way of presenting the brand name. It seems highly probable, however, that it would be most desirable from the point of view of the gasoline company if the symbol became associated with its particular brand and, further, if that symbol tended to be positively associated with some well thought of attributes of gasoline. To the extent that this becomes true, the public would be inclined to recognize quickly and easily the particular symbol and to think positively of it in reference to a particular brand. This implies, of course, that in this process the public will be aware that the symbol does not stand for other brands of gasoline.

GENERAL OBJECTIVES OF STUDY OF GASOLINE SYMBOLS

To what extent are these symbols associated with particular brands of gasoline and, further, to what extent are particular positive and negative attributes of gasolines associated with them? This is the general question of the meaning of the symbols. The study is based upon the assumption, which assumption is itself tested in the study, that these symbols will indeed have some describable meaning and that they will tend to be associated primarily with the brand of gasoline for which they were designed.

The first part of this assumption—that each symbol will have a describable meaning—is based upon the general psychological observation that recognizable and often-seen symbols of this sort do generally tend to be associated in the minds of the public with some particular attributes. This is accomplished, of course, through advertisements and other public statements made by the company—which statements are normally made in connection with the symbol (whether the company makes a direct reference to the symbol or whether the symbol merely occurs in the same context as positive statements about the gasoline it represents).

The second part of the assumption—that the symbol will be identified with the brand in question—is made on the supposition that people do indeed read advertisements and gas station signs and that they will come to recognize the connection between a symbol seen frequently in these contexts, whether or not the company calls attention directly to the symbol.

The specific goals of the study may be stated as follows: (1) to investigate the extent to which meaning is attributed to the symbols; (2) to investigate the extent of recognition of the symbols of gasoline companies; (3) to investigate the extent to which this recognition implies "gasoline;" (4) to investigate the extent to which this recog-

nitition implies a particular brand; and (5) to describe the image of some symbol as seen in the associations given to it.

THE SYMBOLS USED

For purposes of the study, the symbols used by 10 gasoline companies have been selected for study. These are companies generally known in the Chicago area which maintain gas station facilities there. They are: Mobilgas, Shell, Sinclair, Texaco, Pure Oil, D-X, Philco, Cities Service, Conoco, Standard.

Each of the symbols was reproduced in color upon an approximately 5" x 5" white card and the card mounted upon an acetate-covered black background. In each case the brand name was removed though all other aspects of the symbol were retained.¹ In the upper left hand corner of the black background appeared, in white ink, the number assigned to each symbol. No other mark appeared on either the white card containing the symbol or upon the background. In the specific tests and interviews used (described below) the symbols were presented in rotating order.

THE SUBJECTS STUDIED

The study is based upon 50 subjects from the Chicago city and suburban areas. The selection of the cases is outlined in Table 1.

TABLE 1
DISTRIBUTION OF SUBJECTS BY SEX AND SOCIAL CLASS

	<i>Upper Middle Class</i>	<i>Lower Middle Class</i>	<i>Working Class</i>	<i>Total</i>
Male	8	19	11	38
Female	4	5	3	12
Total	12	24	14	50

The objective of this particular sample was to represent the regular car-driving public. All subjects are car-drivers, some also car-owners. The purpose was not specifically to investigate differences among the subcells of this population, there generally being too few cases in the smallest cells for such comparisons. Where such differences appear between sexes (based upon 38 males and 12 females) or between

¹ We wish to express our appreciation to the artists of the *Chicago Tribune*, who reproduced these symbols. It is some evidence of the faithfulness of these reproductions that no subject complained of the art work and none felt that they did not truly represent some real symbol.

social class groups (based upon 12 Upper Middle, 24 Lower Middle and 14 Working Class), they will be noted in the report.

The larger percentage of males was reflective of the higher percentage of male drivers. The higher percentage of Lower Middle Class subjects was generally reflective of the higher percentage of such persons in the population, though no effort was made to make the smallest subgroups specifically representative of the proportions of such persons in the general population. The subjects were interviewed by skilled interviewers either in their homes (predominantly true with the women subjects) or in their offices (somewhat more generally true of the Upper Middle Class male subjects).

TESTS USED

The means by which the various purposes of the investigation were explored are outlined in the following paragraphs. As will be seen, some of the questions ask for entirely free and unstructured comments by the subject, in which the skill of the interviewer to encourage and follow up remarks is stressed. On the other hand, some of the instruments have been designed on the basis of the specific hypotheses and devised in test form. In these, the interviewer's skill is utilized primarily in securing the right mind-set on the part of the subject and in encouraging him to follow the directions.

Test 1: *Free Association to Symbols.* This item is an entirely free association test. In it we present to the subjects each of the symbols in order and request them to tell us the first thing that comes to their mind, what the symbol reminds them of, what ideas, feelings, thoughts strike them as relevant for the symbol. It will be noted that we have not informed our subjects that these are gasoline symbols.

The symbols, while listed in the same order on the recording form, were presented to each subject in a different order. Thus the first subject received the symbols in order of number, 1, then 2, and 3, and so on. The second subject first saw symbol 2, then 3, then 4 and so on, with 1 being the last seen. Subject 3 would first see 3, then 4, and so forth. In this manner we guard against any particular preferences for order of presentation and against the subject becoming either more responsive at the beginning or less responsive at the end of the presentation of the symbols. We are thus sure that the attributes connected with each symbol are genuinely responsive to that symbol rather than merely to the order in which the symbols were presented.

It was anticipated that this test would not produce very rich material. Subjects usually find it difficult to respond very fully under such gen-

eral instructions. We wanted to be sure, however, that they had every opportunity to give as much spontaneous material as possible.

Test 2: Guided Associations to Symbols. This item is designed to assist the respondent in suggesting attributes relevant to the symbols by providing a list representing a fairly wide range of attributes. The subject still has free choice, however, to indicate which of the attributes applies to which symbol. The manner of presentation was to spread the symbols out in front of the subject all at once and in a random order each time. The subject thus looked at the attribute listed on the sheet of paper (a copy of which he was provided to look at or hold in his hand) and then was asked to look over all of the symbols. He was then to indicate which symbol was most appropriately described by each of the words listed and which least appropriately described. The randomization principle in this item is found, of course, in the fact that the symbols were first spread out all at once and in a different arrangement each time, and the subject could pick whichever symbol he chose as most or least appropriately described by the attribute listed. The interviewer then merely recorded the number of the symbol in each instance, plus such spontaneous remarks as the subject made while making his choice.

The words or attributes listed on the recording sheet contain both negative and positive words. Further, they were chosen to represent certain general hypotheses about the meaning of these symbols. Such choice according to hypotheses, as opposed to some entirely random choice of words, permits us to test these hypotheses in terms of the ways in which the symbols were chosen to be associated with not merely individual words, but groups of related words. Grouped according to these hypotheses, the words are:

Friendly	Powerful	Dependable	Scientific	Clean
Kindly	Strong	Confident	Modern	Wholesome
Helpful	Energy	Safe	Efficient	Neat
Impersonal	Sluggish	Weak	Old Fashioned	Messy

These groups may be called, roughly, the Friendly, Powerful, Dependable, Scientific, and Clean groups. They represent five general hypotheses to constellations of attitudes having high probability of being associated with gasoline. They also represent constellations of ideas to be found in gasoline advertising.

The negative items in each of these constellations are presented in large part to permit subjects to indicate the extent to which a particular symbol does not partake of the qualities of one of these constellations.

Test 3: Story Association Test. As a further means of exploring the meaning which subjects attach to these symbols, we have constructed a series of small stories. Each of them calls for the subject to associate a symbol with a particular constellation of ideas as represented by the stories. These are stories of the Boy Scout team, the Sports team, the Yacht situation, the Sorting situation and the Community Chest situation. As will be seen, the first, second, third, and fifth of these situations represent idea constellations. The fourth, the Neutral-Sorting, is included specifically to see to what extent certain symbols are not at all associated with such idea-groups, but are rather seen as entirely neutral symbols that could mean anything. (It will be shown that this is true, that certain of the symbols have no inherent meaning, but are thought of as exclusively neutral. This implies that they could be used by anybody to stand for most anything.)

In this test, the individuals are in the situation of having to select a flag or emblem for each of the situations proposed, and we then ask our subjects to select a symbol from our list that would be fitting for that situation. Again, the symbols are arranged randomly in front of the subject as each story is read to him and he may select any symbol in any order. Following this, his reasons are sought for the selection—thus giving some further near-spontaneous material on the meaning of the symbol.

*Test 4: Brand-Relevance Test.*² Here, we move more specifically to the question of whether or not these symbols are thought to be good representatives of brands or products, and what these might be. Our interest is primarily in the extent to which gasoline is thought of spontaneously as one of the centrally relevant products for which these symbols are appropriate. We do not concern ourselves with specific attributes of the symbols, but with the question of their relevance for products and brands. The symbols were presented to the subjects in rotating order, as with Test 1, and each time they were asked to indicate a product for which the symbol might be appropriate.

Test 5: Symbol Preference Test. We are now testing specifically the aesthetic preference of the subjects for these symbols. It should be recognized that to some extent the preference may be a reflection of the product which they think it might represent. We are, however, introducing no notions of specific brand here, but merely asking for

² This test is omitted in subsequent analysis for the reason that the responses to it scatter too widely.

preference. We secure this by asking for the three best liked and the three least liked symbols.

Test 6: Symbol Recognition Test. It is always possible that these particular symbols are so well known that prior material on the meaning of the symbols would best be thought of as being direct reflections of the subject's attitudes toward the gasolines represented. Here we attempt to test the extent to which this may be true by inquiring as to their recognition of these symbols. We specifically ask if they recognize any of these as symbols for some product (still not mentioning gasoline) and ask for the name of the product.

Test 7: Social Data Sheet. This form is for our interviewers to fill out. It serves merely to check the extent to which the interviewers have appropriately identified our subjects and to check their estimate of social class placement. Since we checked these closely while the interviews were coming in, and hence could make adjustments and corrections, we find that the ideal sample stated under Subjects Studied above was met exactly by the actual subjects chosen for interview.

THE EXTENT TO WHICH MEANING IS ATTRIBUTED TO THE SYMBOLS

Our first general purpose in the study of these symbols is reflected in the results of the first test. As was pointed out, it was not expected that very notable associations would be produced. This is a difficult task for most people. Yet it appears important to ascertain the extent and kind of spontaneous association before providing any guidance to this association, as we do in later items.

First, let us ask the general questions: (1) To what extent are associations produced, of whatever nature? (2) Of these associations, how many may be thought to be associations implying some specific grouping of ideas around the symbols? That is, to what extent are the associations merely reflections of the form, or color, or shape of the symbols, or to what extent do they imply that people hold a more or less clear grouping of specific ideas about the symbol. If people appear to associate in some systematic way with ideas of specific attributes, we might conclude that the symbol has a definable meaning, that people have some organized past experience with the symbol. Whether or not these possible meanings are relevant to the issue of gasoline can be determined subsequently—in other analyses. On the other hand, if the associations are primarily “klang” associations, we might conclude that they are produced merely by the immediate stimulus of the particular symbol and do not represent any substantial previous grouping of meaningful ideas. These “klang” associations are such as the associa-

tion of "white" to the stimulus "blank," or the association "round like a ball" to the stimulus of a circle on a piece of paper. On the other hand, the association "afraid" to the stimulus of the word "black," or the association "gentle, like mother" to the circle-stimulus, might be thought of as associations implying previous experience of an organized sort.

We will call these "klang" associations Form and Color associations. In tabulating them below, we will also include the various miscellaneous associations that do not occur frequently enough to be properly located in the other tabulation category below.

We will call the associations that imply some real meaning of a systematic sort Content and Feeling associations. These two broad categories of response to Test 1 for each of the ten symbols are tabulated in Table 2.

TABLE 2
FREE ASSOCIATIONS TO SYMBOLS

	1	2	3	4	5	6	7	8	9	10
Form and Color										
Number	39	72	89	70	67	75	73	70	56	12
Content and Feeling										
Number	36	13	17	18	12	13	22	13	15	49
Per cent	48	19	19	25	18	15	23	15	21	80
Total Responses	75	85	106	88	79	88	95	83	71	61

In Table 2, it will be understood that the Form and Color associations represent the number of total references in this category made in the free responses of our subjects. Since frequently more than one reference was made by each subject, the totals will always be greater than the number of subjects. Similarly, those listed under Content and Feeling are the total in this category. The section marked Total Responses is the summation of these two. Under the category, Content and Feeling, we have transformed the specific number into percentages of total responses for reasons that will be explained later.

If we look first at the row marked Total Responses, it will be seen that our subjects do indeed have a significant number of responses, producing in all cases more than one idea in response to the stimulus symbol. It tells us little more than this, however, unless we break these responses down into some meaningful categories.

Looking at the row marked Form and Color responses, it becomes apparent that, with two exceptions, the number of responses in this category is very high. Ignoring for a moment those two exceptions

(symbols 1 and 10), it would appear that the greatest weight of responses to the free association item is primarily determined by the immediate qualities of the stimulus—its shape, form, color. The corresponding numbers under Content and Feeling present the inverse of this picture. Here it will be seen that, with the exception again of symbols 1 and 10, the number (and corresponding percentages) of responses occurring in the Content and Feeling category are low—ranging from 15 per cent of the total to 25 per cent.

The implications of these responses are fairly clear: The associations of our subjects to the symbols of gasoline companies are primarily determined by the immediate response to the form and color of the stimulus presented to them in this specific test, and correspondingly that the associations are very limited that imply some systematic recalled meaning and feeling about these symbols.

The two exceptions to these generalizations should now be examined. As will be noted, symbol number 1 (Mobilgas) received 48 per cent of its associations in the Content and Feeling category and symbol number 10 (Standard) received 80 per cent of its associations in the Content and Feeling group. Further analysis of these two symbols is warranted by these heavy percentages, implying a strong attribution of organized meaning.

Symbol 1—Mobilgas. When the specific associations are broken down into their component parts, 36 associations appear in the Content and Feeling category. When these are further categorized, we find:

Ideas of power and speed	29
Positive feeling attributes	5
Negative feeling attributes	<u>2</u>
Total	36

The specific items of positive feeling attributes are those of "interesting, good movement, orderly, light." While these have no very clearly defined meaning, they do imply some thought beyond the immediate stimulus of a positive sort and hence they are included here. The two negative responses are "bad fit" and "running away" and, while negative in tone, they also imply something beyond purely "klang" responses.

The preponderance of content responses, however, are those listed under the rubric Power and Speed. It is thus apparent that at least 29 of our 50 subjects respond to the Mobilgas symbol with gasoline-relevant concepts that are positive in tone.

Symbol 10—Standard. When the associations to symbol 10 are summarized, it appears that 49, or 80 per cent of them are in the

Content and Feeling category. When these are further categorized, we find:

Patriotism theme	31
Related positive theme	11
Other positive attributes	6
Negative attribute	<u>1</u>
Total	49

The Patriotism theme is a highly positive series of references implying a systematized series of attitudes composed of such items as "reminds me of America," "responsibility, freedom," and the like. While it is apparent that the red-white-and-blue aspect of the stimulus is the part that sets off these associations, nonetheless they exist and are presented over and over in the same positive fashion.

The Related Positive theme is essentially the theme of the Olympic games—presumably instigated by the torch portion of the symbol. It is positive, and contains such items as "reminds me of athletic contests, Olympics," "fortitude," "torch lighting the world," and so forth. Even among the miscellaneous attributes, categorized under the Form and Color category and hence not referred to above, there are still some positive, though scattered, references. Specifically there are four—"good torchlight," "a fraternal order," "the medical profession," and "candlelight."

Thus Symbol 10 is responded to in a highly positive fashion by associations systematic and meaningful. There may be some question, however, as to whether these particular associations are what we might call gasoline-relevant. In comparison, the power and speed associations to Mobilgas are quite apparently so relevant. As will be seen by subsequent analyses, however, we are entirely correct in assuming that this general positive aura of high moral quality and good-will attributed to Symbol 10 will be found to be substantiated in other material and reflected in other clearly gasoline-relevant decisions made by our subjects.

In summary, it would appear that our subjects attribute very little spontaneous meaning to these gasoline symbols, responding primarily in terms of immediate reaction to the form and color of the stimulus presented in the test itself. Very seldom do they appear to bring to the interview systematic recollections of meaningful associations. The exceptions appear to be the Mobilgas and the Standard symbols, to which highly consistent, positive, and gasoline-relevant associations are made.

SYMBOL-RECOGNITION AND THE ATTRIBUTION OF MEANING

We now wish to inquire as to whether the particular pattern of low-meaning attribution seen for the majority of symbols in the pre-

ceding section may be thought to be a direct reflection of the extent to which our subjects recall and recognize the symbols used. If this were the case, we would, of course, expect that the symbols 1 and 10, Mobilgas and Standard, would be the most highly recognized symbols and the others would receive very low recognition. A test of this possibility is provided by Test 6, in which we specifically ask our subjects to name the gasoline company or other product which they think belongs to the symbol in question. It is to be noted that Test 6 occurs after five other procedures have passed—after they have had five opportunities to recall the product or brand associated with the symbol. The subjects have not been told what brands the symbols represent, nor that they do indeed represent gasoline. But it still might be presumed that the five repetitions of the symbols would have crystallized more recollection than was present at the beginning of the interview. Thus the amount of recognition must be thought to be the maximum amount of recognition rather than the minimum.

When we now compare the extent of recognition, we find the results shown in Table 3. Two general conclusions of importance can be

TABLE 3
SYMBOL-RECOGNITION

	<i>Recognition Responses</i>	<i>Nonrecognition</i>	
		<i>Responses</i>	<i>No Response</i>
Symbols Most Recognized			
Mobilgas	38	4	8
Shell	41	4	5
Phillips	40	10	0
Standard	34	9	7
Cities Service	28	13	9
Texaco	40	7	3
Symbols Least Recognized			
Pure	17	28	5
D-X	4	39	7
Conoco	12	30	8
Sinclair	4	38	8

drawn from these figures. First, it is apparent that these symbols differ markedly in the extent to which they are recognized by our subjects, even after seeing and talking about them six times. There is a group of highly recognized symbols—Mobilgas, Shell, Phillips, Standard, Cities Service, Texaco; and there is a group of poorly recognized symbols—Pure, D-X, Conoco, Sinclair.

The second conclusion is most relevant for our previous analysis of the extent to which meaning is attributed to the symbols. It may be

seen by further analysis of the group of highly recognized symbols. Recall that only 2 of the 10 symbols were given any significant amount of attributed meaning in Test 1—these were Mobilgas and Standard. However, it does not follow that this meaning is attributed exclusively on the basis of symbol-recognition, since four other symbols are recognized as frequently yet are not assigned a significant meaning. These are Shell, Phillips, Cities Service, and Texaco, where recognition is high and meaning-attribution low.

THE SPECIFIC MEANING AND IMAGERY OF THE SYMBOLS

Up to this point, we have seen that the symbols themselves arouse limited spontaneous associations among our subjects, that in contrast the responses to the unaided associations are determined more generally only by the form and color of the stimulus-symbol itself as presented in our test. Two exceptions were noted, the Mobilgas symbol and the Standard symbol. It has further been pointed out that the greater association to these latter two symbols is not entirely a matter of recognition, since other symbols were similarly recognized. Now we need to examine more specifically the meaning and imagery which appears to characterize these symbols when we provide our subjects with some ideas against which to make judgments and choices about the symbols. Some of the specific imagery that attaches to the Mobilgas symbol and the Standard symbol was discussed above.

It will also be important to see to what extent this same imagery is characteristic of the aided associations. If it is not characteristic of the aided associations, we should re-examine our findings with the thought in mind that the associations gained in the spontaneous Test 1 were random and not truly representative of some systematic body of previous notions held about those symbols. If the picture of the Mobilgas and Standard symbols is considerably different, we would wonder if perhaps our conclusions about the absence of systematic associations to the other symbols were incorrect.

In Test 2, the guided-association item, we asked our subjects to indicate which of the 10 symbols were most, and least, appropriately described by a series of words. The words chosen represented a series of hypotheses about the possible meaning of the symbols. These words were grouped into five hypothesis-groups represented by the key words: Friendly, Powerful, Dependable, Scientific, and Clean. It will be noted that these groupings represented themes notable in advertising about gasoline and that they were ideas that might be thought to be clearly gasoline-relevant.

To explore the meanings of the symbols in terms of these gasoline-relevant attributes, we have tabulated all of the responses to each attribute according to the symbol which was thought appropriate, or not appropriate to it. For purposes of presentation, we will divide the symbols studied into categories of: (1) High Definition, referring to those receiving the highest two categories of responses (Mobilgas and Standard symbols); (2) Medium Definition, referring to those receiving lower numbers of responses but where there is still some distinction from those of low definition (Sinclair, Pure, D-X, and Conoco); and (3) Low Definition, referring to those receiving either very few mentions on attributes or where the votes did not distinguish the symbol from many others on individual attributes (Phillips, Cities Service, Shell, and Texaco). As will become apparent, both of the High Definition gasoline symbols are, in addition, positively perceived. This is not true of the Medium Definition symbols, which contain both negatively and positively described symbols. The Low Definition symbols are generally negatively perceived.

High Definition Symbols. The two symbols called High Definition symbols are those for Mobilgas and Standard. This is consistent with the extent to which these two symbols were given prominence in the free association item discussed earlier. Considerable refinement on the definitions available from that item is possible from the guided-association material. The items on which these two symbols received notable mention are summarized in Table 4.

Examination of Table 4 will quickly reveal the picture with regard to these two symbols. The table is to be understood to be a list of the extent to which our subjects selected the symbols in question as best described by, or most appropriately representing, the attribute listed at the left. In each case the numbers listed are the highest and second highest selections for the attribute.

The Standard symbol is by far the most clearly defined and by far the most positive symbol. It is seen by our subjects as: Friendly, Confident, Efficient, Neat, Kindly, Dependable, Wholesome, Helpful, Safe, Strong, Scientific, Clean. The items, Scientific and Clean, are last in importance. This is true since the number of mentions, while second highest of all symbols, are nonetheless rather low. They are mentioned for Standard, but not for Mobilgas since the Mobilgas mentions on this item are only 3 and 1.

The other side of this coin may be seen by examining Part B of Table 4, which lists the attributes thought of as least appropriate for a given symbol. It may be seen that the Standard symbol receives

TABLE 4
MEANINGS ATTRIBUTED TO GASOLINE SYMBOLS

	<i>High Definition</i>		<i>Medium Definition</i>				<i>Low Definition</i>		
	<i>Mobil-gas</i>	<i>Standard</i>	<i>Sinclair</i>	<i>Pure</i>	<i>D-X</i>	<i>Conoco</i>	<i>Phillips</i>	<i>Cities Service</i>	<i>Texaco</i>
A. Most Appropriate									
Attributes									
Friendly	10	19							
Confident		20							
Weak				9	13			10	
Energy	30								
Efficient		14				12			
Messy			14					9	
Neat		13		9					
Kindly		15							
Old fashioned				12			8		
Dependable		21							
Powerful	31								
Modern						22			
Sluggish			11				7	8	7
Wholesome		17		10					
Helpful		19							
Safe		14	8				7		
Impersonal				9	12		7		
Strong	18	16							
Scientific		9				13			
Clean		7		15					7
B. Least Appropriate									
Attributes									
Friendly			9				8		
Confident			12		9				
Weak	10	15							
Energy			11	12					
Efficient					11			10	
Messy					8				8
Neat			15					9	
Kindly			12		8				
Old fashioned		8				15			
Dependable			14	9				9	
Powerful				11	11			10	
Modern			8		10		8		
Sluggish	19			7					
Wholesome			13		7		7		
Helpful			12		9				
Safe			14		11				
Impersonal	10	7							
Strong			9					12	
Scientific			11	9					
Clean			16				10		

few negative votes, even for least appropriate items, but what it does receive are: not Weak, not Old Fashioned, not Impersonal.

It would appear that, in essence, the Standard symbol stands in the minds of our subjects for any good item that one cares to mention.

It is of some interest, however, that this highly received and positively evaluated symbol is not notable on several items—specifically, it is not Powerful, nor Modern, nor high on Energy. While it is mentioned, it is hardly outstanding even in being Scientific and Clean. In fact, it would be appropriate to propose that the high positive definition of the Standard symbol is a blend of friendly, stable, dependable, helpful concepts. While it is agreed that it is efficient and strong, these are less notable elements of its meaning and imagery for our subjects.

The Mobilgas symbol is the second most clearly defined and most positively received symbol. It is seen by our subjects notable on the items: Energy, Powerful, Strong, Friendly.

The Least Appropriate attributes for the Mobilgas symbol, seen in Part B of Table 4, help to confirm this picture. As seen there, Mobilgas is notable in being: not Weak, not Sluggish, and not Impersonal. In this connection, the high votes for Mobilgas on the not Impersonal item are interesting. They serve to confirm the secondary but apparently important element of Friendliness that accompanies the imagery of power and strength.

It is apparent from these items that the imagery of Mobilgas, in some contrast to the imagery of Standard, is one of strength and power, with friendliness. This observation is important for two reasons. First, it describes the imagery of Mobilgas as contrasted with the Standard symbol. Second, it points clearly to the fact that high positive recognition and meaning attribution for these symbols is not random. That is to say, the highly recognized symbols do not receive the high mentions, as was pointed out earlier, and further the highly defined symbols are so defined for quite different reasons.

Medium Definition Symbols. The medium definition symbols, those of Sinclair, Pure, D-X, and Conoco, are those where a less clear-cut imagery exists (as compared with other symbols) or where the imagery that does exist is of a lower order of note.

The imagery of the Sinclair symbol is predominately negative. It is, in fact, the most thoroughly derogated symbol in the group, including the symbols of lower definition. For our subjects, the most appropriate attributes of the Sinclair symbol are: Messy, Sluggish, Safe. The item Safe is worth special note here. It might be thought to be a positive item in an otherwise negative picture. We might say, for example, that in spite of being messy and sluggish, it is at least safe. Examination of Part B of Table 4, however, reveals that the Sinclair symbol receives an even higher number of votes, 14, for not Safe. It would thus appear, that people are at least in some doubt as to this quality of the

Sinclair symbol. Review of the other Least Appropriate items shows that Sinclair is even more systematically derogated than appeared in the positive choices. There it receives notable mentions for: not Friendly, not Confident, not Energy, not Neat, not Kindly, not Dependable, not Wholesome, not Helpful, not Scientific, not Clean, not Modern, not Strong. In many ways the Sinclair symbol is the reverse of the Standard symbol. Just as there appeared to be few positive items people were not willing to attribute to Standard, here there appear to be few negative items that do not go with Sinclair.

The imagery of the Pure symbol is essentially one of old-fashioned cleanliness. While this is not necessarily a negative image, it is one that bears little relationship to the advertised virtues of gasoline. Thus, the Pure symbol is notable on the items: Old Fashioned, Clean, Wholesome, Impersonal, Neat, Weak. The mentions on Old Fashioned and Clean are the highest among all of the symbols, the others being in each case the second highest. Correspondingly, the high mentions for Pure on the Least Appropriate items are: not Energy, not Dependable, not Powerful, not Scientific, not Sluggish.

The item "not Sluggish" is included because it meets the technical qualifications of being the second highest mention. However, the highest mention is 19 for the Mobilgas symbol. It is probable that we should not even include this low mention of 7 for Pure, though perhaps it is part of the general imagery of Pure as fairly positive but still not really powerful or strong or energetic. The statement that Pure is not Sluggish is also in part contradicted by the fact that these subjects give 5 mentions to the accusation Sluggish in the Most Appropriate items. In contrast, the item Weak receives the second highest votes on Most Appropriate for Pure and only 1 vote for Weak as Least Appropriate for Pure.

The imagery of the D-X symbol is predominately negative and weak. In the Most Appropriate attributes, D-X receives the highest mentions of any symbol on the items Weak and Impersonal. Some reflection of the generally weak nature of the D-X symbol imagery is found in the fact that D-X is notable on no other attributes among the Most Appropriate items. The weak impersonality of the D-X symbol is substantiated by a review of the items thought least appropriate for D-X. They are: not Confident, not Efficient, not Messy, not Kindly, not Powerful, not Modern, not Helpful, not Safe.

This picture of Least Appropriate items is interesting for three special reasons. First, it documents the imagery of weak—in the high votes on not Powerful, not Modern, not Efficient. Secondly, it docu-

ments the imagery of impersonal—in the high mentions on not Kindly, not Helpful, not Safe. The note on not Messy is a third point of interest. It would appear to refer primarily to the symbol itself, as a clear, simple diamond, and not to any imagery of D-X as the opposite of Messy, Efficient, Modern, Strong, Scientific.

The imagery of the Conoco symbol is essentially one of efficient modernity, unaccompanied by any other positive gasoline-relevant attributes. The imagery of the Conoco symbol is defined by its high mentions on the Most Appropriate items as follows: Modern, Scientific, Efficient.

It is of interest that this imagery of efficient modernity constitutes the only part of the meaning of Conoco. Conoco is notable on no other positive or negative items among the Most Appropriate. It is, however, almost the only gasoline symbol with any mention of Modern—the next highest mention on this item being one of only 6 mentions.

On the Least Appropriate items, Conoco is the highest mentioned on the item not Old Fashioned. This item, the only one of any note among the Conoco least appropriate items, merely substantiates the view of Conoco as Modern and further serves to show how delimited the Conoco imagery is to this particular view—all mentions on other items being negligible.

Low Definition Symbols. The symbols of low definition, Shell, Texaco, Phillips, and Cities Service, are those where the number of mentions on specific items are low and where the pattern of mentions presents no clear-cut imagery. They are either negative in tone or merely indifferent in meaning.

The imagery of the Shell symbol is notably weak and composed of poorly defined elements of cleanliness and sluggishness. Thus, the two items on which our subjects note the Shell symbol on the Most Appropriate attributes list are: Sluggish, Clean.

Even here, however, the votes are only 7, with the symbols that receive firmer definition receiving 11 and 15 votes respectively on these items. When we examine the Least Appropriate items, we might well expect, as with Sinclair, to find a more firmly defined negative picture. This is not the case with the Shell symbol, there being no single item worthy of note. It is true that Shell receives 7 votes for not Scientific, but in an array of mentions which makes this number third in size and hence not important. This imagery of the Shell symbol suggests more a general aura of lack of feeling and lack of clear-cut meaning. Our subjects do feel it is probably sluggish, though this feeling is

not marked, and yet it is probably also clean, though this too is not marked. It should be recalled that this absence of definition is not due to failure to recognize the symbol. The Shell symbol received, in fact, the highest number of recognition responses (41) of all the symbols, being even higher than the two symbols receiving the most clear-cut definition (Mobilgas with 38 recognition responses and Standard with 34).

The imagery of the Texaco symbol is the weakest of all symbols, receiving no outstanding mention of any kind. On the Most Appropriate attribute list, Texaco receives no vote that is at all outstanding. It is true that it receives 7 mentions on the item Friendly, though in a situation in which the next highest mention is for 15 votes. It also receives 6 mentions for Clean, but this number of votes is the fourth from the highest and hardly defines any significant meaning for the symbol. These two minimal mentions, Friendly and Clean, are indeed positive ones, however, and might be thought to reflect some weak but positive imagery. It should be noted, however, that Texaco also receives 6 votes for not Friendly and 7 for not Clean. These certainly cancel out any positive gain from the Most Appropriate items and leave the Texaco symbol as one with no definable meaning.

The one item which might be at all worth note is a mention of 8 votes on the item not Messy among the Least Appropriate items. This is the highest vote on this item. Were the number higher, one might think that the symbol at least is thought to be fairly neat in appearance. Yet the votes on the Messy item for other symbols scatter in a meaningless fashion; 8 is the highest and there are votes of 8, 7 and 6 for other symbols. In addition, Texaco receives only a single vote of Neat among the Most Appropriate items.

It should also be noted that the Texaco symbol is among those receiving high recognition. Thus, as noted earlier, Texaco receives 40 recognition responses, higher than either of the two most clearly defined symbols. It is not that Texaco is not recognized, but that it carries no definable meaning.

The Phillips symbol is viewed with a constellation of items reflecting a weak definition of old-fashioned impersonal safety. It receives only moderately high votes, no outstanding ones, on the Most Appropriate items: Old Fashioned, Sluggish, Safe, and Impersonal. Correspondingly, it receives mentions on the Least Appropriate items: not Friendly, not Clean, not Modern, and not Wholesome.

This is an image essentially negative in tone, old-fashioned, sluggish, impersonal though safe. Even the negative votes present no particularly

clear-cut definition, merely attesting to the lack of any real body of systematic feeling about the symbol.

Again, it may be noted that Phillips is one of the high recognition symbols. It receives 40 recognition responses. While the symbol adequately calls to mind Phillips gasoline, it reflects only an image of a kind of impersonal sluggishness.

The imagery of the Cities Service symbol is one of weak sluggishness. Cities Service symbol receives mentions on the Most Appropriate items: Weak, Sluggish, Messy. The corresponding picture from the Least Appropriate items extends somewhat this imagery. It receives notable mentions on: not Efficient, not Neat, not Powerful, not Strong, not Dependable. Thus the negative items agree with the image of weakness and sluggishness and in fact specifically underline the sluggish aspect in the clear rejection of the attributes Powerful and Dependable as having anything to do with it.

Cities Service is the lowest of the High Recognition symbols, receiving 28 mentions. This is still considerably higher, however, than the highest of the low recognition symbols—Pure, which received 17 mentions. Thus again the imagery of Cities Service does not stem from failure to recognize the brand it represents.

To this point, we have provided a beginning view of the imagery of each of these gasoline symbols in terms of the various positive and negative attributes which our subjects attribute to them. It is apparent that each of these symbols does indeed have a public reputation. This public reputation differs in two ways—in the extent to which it is strong and weak, that is, in the extent to which the subjects hold a clearly defined versus a diffusively perceived image of the symbol. They also differ in the particular content of this imagery. Some of these specific meanings have been described as they are available to us from the free- and from the guided-association material. It is also further apparent that these images are not dependent upon the extent to which the symbols are recognized by the subjects as symbols for particular gasoline companies. Further material is available from the other tests. This material will serve to elaborate and extend the images and reputations presented so far.

THE CONTEXTS SEEN AS APPROPRIATE TO THE SYMBOLS AND SOME FURTHER MATERIAL ON THE REPUTATION OF THE SYMBOL

A further effort to explore the imagery and reputation of these symbols is to be found in the story situations presented to our subjects, after which they were asked to indicate which of the symbols would

appear to them as most appropriate for those contexts. These contexts were the Boy Scouts story, the Sports Club story, the Yacht story, the Sorting story, and the Community Chest story. It will be noticed that each of the story contexts is aimed at a little different imagery about public life.

Thus, the Boy Scout story might be thought to elicit either notions of protective care or possibly the attributes of high moral character. We would be inclined to assume that symbols which are in any substantial proportion thought to be best to represent this story would be ones in which either of these two attribute elements were high in the minds of our subjects.

The Sports Club situation is one in which the active and energetic elements might be thought to be high, and hence the symbols associated with it might be reflective of these attributes thought to be in the symbols themselves.

The Yacht story is a little more diffuse than the two above, but it was thought it might represent either some notion of luxury, or of active, outdoor, carefree attributes.

The Sorting story was designed as a neutral situation, one in which no particularly meaningful attributes were imbedded. It was thus proposed that the symbols associated with it, as opposed to the other stories, might be those symbols most neutral and meaningless to our subjects.

The Community Chest story was thought to be one of public responsibility as well as one emphasizing some of the high moral character elements.

Two kinds of data were available on these story situations. First, we asked our subjects to indicate the symbol they thought best for a particular situation. Second, we asked them their reasons for this choice, in the form of the questions about what meaning they thought the symbol should have for the people in the situation. This second element provided further data on the meaning of the symbol and a specific check on whether or not the assumptions made about the symbols hold for the subjects.

Symbols Appropriate for Each Situation. In summary, we find the following choices of symbols for each situation:

The Boy Scout situation is best represented by the Standard symbol. Of the 50 subjects, 18 chose this symbol. The other 32 choices scatter meaninglessly throughout all other symbols, the next highest choice being only 7 mentions.

The Sports Club situation receives two high mentions, those for the Mobilgas symbol and the Sinclair symbol.

The Yacht situation also received two high mentions, those for the Mobilgas symbol and the Texaco symbol.

The Sorting situation is the only situation for which four symbols were chosen in about equal amounts, these four all being notably higher in choice than the votes given to the other symbols. These are the symbols for Sinclair, Texaco, Pure, and D-X.

The Community Chest situation receives two high mentions, those for the Standard symbol and the Texaco symbol.

It is thus apparent that the symbols for Phillips, Conoco, and Cities Service are chosen for none of the situations in any significant proportion.

Meanings Attributed to Symbols for Each Situation. Now let us examine the specific meanings attributed to these symbols. This will permit us to test our assumptions about the reasons for the choice of symbol for situation, and to extend the information on the reputation of the symbols themselves.

Boy Scout Situation—Standard Symbol. When we asked our subjects what meaning they thought the Standard symbol should have for the boys in the scout team, they unanimously (all 18) agreed on the Patriotism theme as most appropriate. Statements of the theme included such individual items as: meaning of the flag and freedom; purity, courage, and freedom; pride, justice, liberty, hard work; courteousness, helpfulness.

It may be noticed in passing that this pattern is closely analogous to the pattern seen from the guided-association item. It becomes increasingly clear, therefore, that the high positive evaluation placed upon the Standard symbol is that of high moral character, sound, good, solid American virtue.

Sports Situation—Mobilgas and Sinclair Symbols. The responses to the question of meaning of these two symbols make it clear that they are associated with the Sports situation for two entirely different reasons.

First, the Mobilgas symbol. Here all 13 subjects who chose this symbol unanimously agree that it represents the constellation of Speed-Power-Action. These are the specific words in which the subjects express this theme. In addition, there are a few scattered responses given by the same subjects who gave the Speed-Power-Action responses, such as "teamwork," "coordination," "clean sportsmanship."

It may thus be seen that the imagery of the Mobilgas symbol in this second major test closely parallels the imagery of Friendly, Powerful, Energy, Strong seen in the guided association element. To this con-

stellation, the situations test would add the closely related item, Action.

Second, the Sinclair symbol. While the Mobilgas symbol receives the highest votes for this Sports situation, the Sinclair symbol is second in importance and significantly higher in mention than the remaining possible symbols. There are 9 such choices for the Sinclair symbol. Of these 9, 8 are essentially "klang" associations. Their content is something as follows: "the symbol is a target or a ball and these have to do with sports." It is thus apparent that this choice is based on no direct reputation of action, or other positive attribute association with Sinclair. It is a direct association to the form and color of the symbol as presented in the stimulus situation—it does indeed look like a ball or target, as our subjects so generally claimed in the first free-association item.

Yacht Situation—Mobilgas and Texaco Symbols. In the Yacht situation, there are two high items, the Mobilgas and the Texaco symbols. All others receive negligible mention.

First, the Mobilgas symbol. The associations to the question of what this symbol might mean to the persons in the yachting situation further elaborates the Mobilgas imagery. The distribution of responses may be categorized:

Speed-Power-Action	6
Carefree flying	5
Pleasant emblem	2

In other words, the Speed-Power-Action complex is again apparent. This time, however, a slightly altered note is added, that of the carefree element. This element clearly stems from the "flying" horse and comments relate this to the skimming, flying over the water of the yacht. It is of moment that this carefree flying element does not appear in other contexts, even in other story situations. (It is quite true that the horse is recognized as a flying horse in other situations, but the flying-carefreeness element is not previously noted.)

Second, the Texaco symbol. The use of the Texaco symbol here is closely analogous to the use of the Sinclair symbol in the Sports situation. That is to say, its use appears to reflect no particular positive attributes that are gasoline-relevant, but rather some attributes, some positive, that stem directly from the stimulus qualities of the symbol itself as presented here. The responses to the question of what meaning the symbol would have for the yachters may be summarized:

Star equals sailing	6
Symbol easy to see	4
Star equals luxury	5

All of these groupings are essentially "klang" associations of the order of the "ball equals sports" type seen in the Sinclair symbol. Here the star is the focal point. In one case, they propose that stars equal night and that would be pleasant sailing. A few utilize the concept of a star-boat. The second grouping is essentially a meaningless one, referring merely to the notion that the star would be easy to see and hence one could recognize your boat better. The third grouping is an association to the concept of the star as representing something of quality, of luxury, of first place. It represents the extent to which the stimulus of the Yacht situation does indeed arouse notions of luxury. These associations might be thought to be somewhat positive, hardly markedly so. This interpretation is consistent with the picture seen in the guided-association item—at least to the extent that there is no clear-cut definable imagery for the Texaco symbol, but at least also no clear-cut negative attributes.

The Sorting Situation—Symbols for Sinclair, Texaco, D-X, and Pure. One theory at the outset was that the symbols of the ten brands might be roughly arranged in the order in which it was probable that meaning could be attributed to them. Among those that seemed essentially neutral in this analysis were those that appear to have no particular inherent meaning. For these, we picked the Sinclair, D-X, and Pure symbols. We excluded the Texaco symbol on the logic that the star would arouse the associations we have seen in the response to the Yacht situation above. We also proposed that these symbols would be highly chosen in the sorting situation, on the logic that this sorting task was essentially a neutral one with no specific meaning. As is seen, the symbols for Sinclair, D-X, Pure, and Texaco are the ones most highly selected for this task.

All of the associations to the question of why these symbols were selected for the sorting task, and what meaning they would have, are "klang" associations. Thus for Sinclair (recall that the task was one of sorting apples):

Red equals round equals apple	4
Target equals noticeable	7

Thus, here the subjects propose that if you're sorting apples, the red of the Sinclair symbol, or its roundness, would be appropriate. If not that, then the fact that it looks like a target would be a good sign for indicating into which box something might go.

The Texaco symbol was also one selected for the sorting test. The logic appears to be that of the prominence of the star, or its representativeness of quality. It is of considerable note that this version of the star

as quality appears only in the situations test, the one fairly well removed from gasoline. In other situations, no such potentially positive imagery appears. Here, however, the responses are as follows:

Star equals quality equals good apples	4
Star is noticeable	5

It might thus appear that, while these are still clearly purely "klang" associations, they are of somewhat more positive quality than others.

For the D-X symbol, the responses are as follows:

Noticeable target	5
Don't know	2

These responses are, in comparison to the Texaco ones, clearly more meaningless and neutral. The symbol can be seen, and presumably anything that stands out would make a good marker for an apple sorter. The two "don't know" responses presumably attest to the difficulty of these subjects in verbalizing their logic.

For the Pure symbol, the responses are as follows:

Neat, orderly, recognizable	11
Red equals apple	2

In this group of four symbols, the Texaco symbol is the one that does have some meaningful background, though low in intensity and seldom connected with gasoline-relevant attributes either here or in other items.

The Community Chest Situation—Standard and Texaco Symbols. The Community Chest situation was one in which elements of public responsibility and/or high moral character attributes were thought to be relevant. This seems sustained by the choice of symbols and their meanings. Thus, the symbol chosen by far the most frequently is the Standard symbol, with 22 first choices. The second is the Texaco symbol, with 11 first choices.

First, the Standard symbol. For the Standard symbol, the associations fairly well repeat the patriotism theme seen in the free-association item and in the responses to this symbol in the Boy Scout situation. Here, the responses are as follows:

Patriotism theme proper	11
Helpful, guiding light	6
Helping others	2
Miscellaneous (dignity, world community, don't know)	3

While we have broken down these elements in the above tabulation, it is apparent that they are all fairly much the patriotism complex seen

elsewhere. Here it is merely that certain elements, particularly the helping element, have been separated out by some people.

Second, the Texaco symbol. The use of the Texaco symbol here is identical with its use on the other story situations, that is, essentially modified "klang" associations to the star. Thus, the responses are:

Noticeable	2
Star equals quality equals high goals	8
Star equals money	1

The last single response appears to be someone who thinks more of the bother of the Community Chest collecting money than of the other aspects more frequently seen by others. The dominant response, however, is the association of the star with notions of quality and good goals.

SOME ADVERTISING IMPLICATIONS

The imagery and reputation of gasoline brands and companies can be described by an analysis of the specific brand symbols used as signs of the company. In general, moreover, the opinions and feelings which consumers hold toward these symbols are more notable and clear-cut than those opinions and feelings which they are able to express toward the brands and companies. This holds true regardless of whether or not the consumer recognizes the symbol as representative of a particular brand.

It thus seems highly probable that there exists, in the attitudinal system of the consumer, a complex of feelings toward gasoline companies that has been minimally utilized in direct public advertising, and, further, that when (as is generally true) the advertising is directed toward the direct technical-product aspects, this complex of attitudes is not appropriately tapped.

It would appear that the clues which the consumer seeks in the effort to decide that a given brand is good or bad for him to purchase are somewhat more inchoate and low in his awareness.

This suggests that a strong program of advertising and public relations that is indirect and evaluative in nature is most apt to reap the advantage of this already present willingness to make purchase judgments on the basis of the interpersonal and moral characteristics of companies. In this direction, the technical aspects of a product should not be ignored—they form the rationale by which the consumer justifies to himself the essentially moral judgment which he makes in choosing a gasoline company. (This might be restated to point out that the consumer does indeed appear to select a gasoline company and not a gasoline.)

The symbols of company identity (referring now only to the signs and symbols studied here) form potentially strong points of identification with a company.

The symbols of the companies of high definition are viewed in quite specific and well-defined ways by the consumer. The symbols of companies low in definition are perceived in a neutral or negative fashion. Some of the symbols appear to be neutrally or negatively perceived through the directly unpleasing or meaningless nature of the symbol. Yet no symbol is perceived in a neutral or negative way if it is the symbol of a highly thought of company. This suggests that, while clearly some symbols are better than others, the symbols can be effectively invested with positive attributes. This could be accomplished through utilizing the symbol in a planned way in a systematic public relations effort.

19. PLAY TECHNIQUES FOR INTERVIEWING ON DURABLE GOODS*

The projective method described in this brief note is in many ways representative of group techniques. It differs from the projective methods presented so far in that data are compiled on the basis of observing respondent actions rather than from verbalization. Unfortunately, it does not differ from other projective techniques with regard to validity, which, as the author noted, remains to be established.

How can you measure attitudes toward durable goods in special classes? What things have influenced the buying patterns in the particular field? What are the things that can be done to influence the present attitudes?

These questions and others have been answered on a small scale by using the "play" techniques, which have been fairly widely used on preschool and nursery school children in finding attitudes. The play tests were made on adult home-owners, who from income and size of lot were prospective purchasers for garden tractors. The sales problem was the broadening of the garden-tractor market (present buyers own

* Adapted from an article by C. W. Garber, Jr., Design and Research Consultant, "Play Techniques for Interviewing on Durable Goods," *Public Opinion Quarterly*, Vol. 15 (Spring, 1951), pp. 139-40.

on an average of three to five acres of property); specific questions were:

How much does the small-home-owner identify himself with garden-tractor advertising?

What does he know factually of garden tractors—uses, attachments, costs, operations?

An interviewing kit was made up for respondents to "play" with, consisting of a play area representing a garden plot, a rake, a hoe, a shovel, two garden tractors. Respondents were given a few minutes to touch the models and become familiar with the models; then the actual interviewing started.

Respondents were asked to prepare seed beds in the play area, using the rake, hoe, and shovel—testing their familiarity with gardening in general. Then they were asked to prepare a new bed using either tractor and any implements; then the other tractor and any implements, or attachments. The interviewer watched for:

General comments about garden tractors or their use.

Remarks relating the use of the tractors to the small-home-owner's property (the respondent).

Familiarity with materials in the kit.

Positive movements—using the models with confidence and sureness.

Hesitant movements.

Rejections of equipment, attachments or method.

Search for objects not included in the kit.

The present (but continuing) study has been small and confined to a single trading area. No tests of validity or reliability have been made nor have there been any control measures paralleling the "play" study. But results have been productive, producing the answers to the stated questions.

Comments and hesitance-rejection have been the most fertile sources of attitudes and convictions. Comments and interest have been the best source of determining "identification-with-advertising." Copy themes and expectation-of-buying patterns will be bonus answers from the study, if the present pattern of answers continues.

There are several reasons why the "play" method is good for this kind of an interview, but the most important is that it is indeed a "projective" technique, delving below the conscious level of response. Second, there has been almost no respondent fatigue; home-owners seem to enjoy the "game" so much that they continue the interview past the play stage to ask questions about gardening. Third, almost no inter-

viewer bias can be introduced, because the instructions do not relate to the answers, but to something that the respondent must do.

The disadvantages are that the interviews are long (have been at least one hour long in all cases) and that there will be a problem in standardizing responses and comparing answers to any other scale of interviews.

Present results, however, justify continuing the series for the advantages seem to outweigh the disadvantages.

20. A SOCIOPSYCHOLOGICAL APPROACH TO CONSUMER BEHAVIOR

The projective and related techniques that comprise such a large part of this volume are based on Freudian psychology that people are motivated in their actions, and presumably in their market behavior, by suppressed fears and desires which can only be brought out by indirect interviewing methods. In contrast to this is the social-psychology approach of Kurt Lewin and others, which teaches that people's actions are survival oriented, their behavior being motivated by their wants and satisfactions in relation to the social environment in which they live.

The present article is a clearly stated example of this latter approach, representing a thought-provoking attempt to explain consumer behavior in terms of imbalance between strength of desire for particular goods and desire to avoid the accompanying expenditures. Particularly interesting is the attempt to relate this rationalization process to economic factors.

In the course of his work with the National Research Council during World War II, Kurt Lewin became interested in consumer behavior. For his analysis he employed the concept that consumer purchasing in its essence involves a psychic conflict between the person's desire for the item in question and his resistance against the undesirable considerations which the purchase entails, such as money cost and buying time involved.¹ He then attempted to infer the magnitude of these tensions for various food items by asking a series of questions and

* Adapted from an article by Warren J. Bilkey, University of Connecticut, "Psychic Tensions and Purchasing Behavior," *Journal of Social Psychology*, Vol. 41 (May, 1955), pp. 247-57.

¹ K. Lewin, "Forces behind Food Habits and Methods of Change," *National Research Council Bulletin*, No. 108, 1943; T. M. Newcomb *et al.* (eds.), *Readings in Social Psychology* (New York: Henry Holt, 1947).

noting which foods were mentioned in answer to each question.² *A priori* his results appear to be very plausible, but Prof. Lewin failed to establish his point because he did not relate these findings to the interviewees' actual purchasing behavior. He merely showed that psychic tensions regarding various food items differed as between high-, medium-, and low-income families, and that within each income group the tensions differed according to the item concerned. The purpose of this study was to determine whether people's psychic tensions regarding selected items relate to their purchases of those items as Lewin postulated.

METHOD

For this study 63 families were obtained, mostly from the Storrs, Connecticut area, who were willing to keep expenditure records for as close to one year as possible and to be interviewed monthly during that time.³ The interviews involved having the family purchasing heads give a numerical rating of their psychic tensions regarding various items on a "0-100 self-rating scale" (see Figure 1).⁴ (The specific questions asked to obtain these tensions varied according to item, and will be indicated as these items are discussed.) Each interview was made as near to the middle of the month as circumstances permitted, and the family's expenditures for the month in question then were recorded during the following month's visit. By this means, the psychic tension data for each month were obtained prior to the completion of that month's expenditures. This permitted an analysis in terms of time sequences, even though a cause and effect relationship was not necessarily established.

To minimize the possibility that the families might deliberately adjust their purchases to harmonize with their interview statements, the following precautions were taken: (1) most of the questions called

² Lewin, *op. cit.*

³ Of the 63 interviewees, 45 were interviewed once a month for 12 months, 8 for 6 to 11 months, and 10 for 1 to 5 months. In terms of occupation: 46 were on the faculty or staff of the University of Connecticut, 14 were small businessmen or laborers, and 3 were married university students. In terms of location: 12 were from New London, Conn. and the remainder were from the Storrs-Willimantic area. No significant differences in tension-purchasing relationships were found between any of these groups.

⁴ This method of measuring psychic tensions was used in preference to Lewin's inferential technique because it avoided some of the major difficulties which the latter entails, viz., (1) the determination of what weights to apply to the interviewees' statements, and (2) the fact that people dislike repeating statements which they feel are either obvious or can be inferred from their earlier answers (see Lewin, *op. cit.*). The latter difficulty is serious because Lewin's method assumes that the frequency of reference to a particular item is an index of its importance to the person in question.

for numerical responses, the full significance of which was not explained to the interviewees, (2) the interviewees never were shown their preceding month's answers, and (3) a large number of questions were asked so as to make memorization of their numerical responses difficult—each interview lasted about 1½ hours and covered all of the family's major disbursement categories. For these reasons it

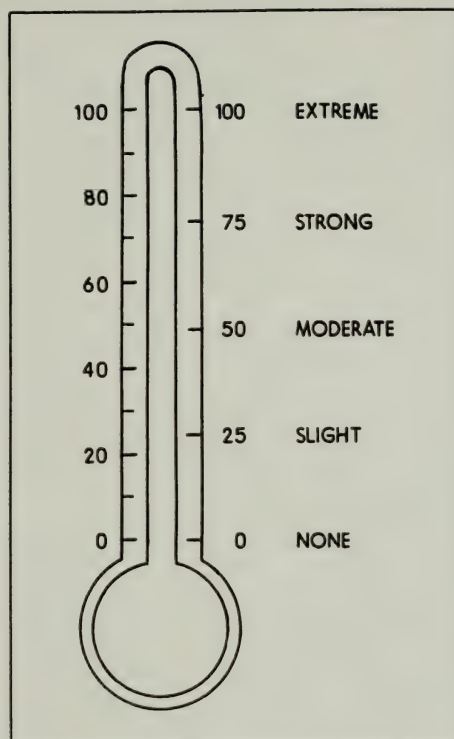


Fig. 1. Self-rating scale (actual length, six inches),

seems unlikely that the families involved could have deliberately adjusted their purchases so as to conform with interview statements, or to have adjusted their responses so as to make their interview statements conform with their anticipated expenditures.

FINDINGS

Consumer Durables. Several of the interviewees indicated that they were giving some consideration to the purchase of a consumer durable such as a car, washing machine, rug, or sewing machine. During each interview they then were asked to give a numerical rating of their psychic tensions for the item in question. Using the thermometer scale shown in Figure 1 they were asked: (1) "How strong is your de-

sire for the (item)?" and (2) "How strong is your desire to avoid the expense which the (item) would entail?" Table 1 illustrates the responses obtained for an item which finally was purchased—a rug. Note the general tendency after June for the desire answers (obtained from question 1 above) to increase relative to the resistance answers (obtained from question 2 above). This indicates an increasing probability of purchase. In every case where a purchase was made, the

TABLE 1
PSYCHIC TENSION RELATIONSHIPS WHICH CULMINATED
IN A PURCHASE—A RUG

(Data as obtained from Interviewee 4; comments made by her during the interviews are given below.)

	<i>Feb.</i>	<i>Mar.</i>	<i>Apr.</i>	<i>May</i>	<i>June</i>	<i>July</i>	<i>Aug.</i>	<i>Sept.</i>	<i>Oct.</i>
Desire for rug	75	—	50	25	25	—	75	90	fin.
Resistance to expenditure for rug	0	—	25	0	100	—	50	50	fin.

Feb.—Had an even stronger desire for a new stove (desire = 100, resistance = 0). The interviewer unfortunately failed to ask amount of resistance against the expenditure of both a rug and a stove together.

Mar.—Bought a new electric range. March interview was missed because of sickness in the family.

Apr.—No comments made.

May—Interviewee commented: "Now that I have a new stove I'm so satisfied that I have no desire for a new rug. Also there's the fact that I got new shoes and a dress." (She had been clothes shopping a few days before the May interview.)

June—Prior to the interview she had priced rugs and found that the kind she wanted would cost around \$300; her earlier estimates had been that such a rug might cost around \$150. She stated, "I simply won't pay that!"

July—Missed interview because family was vacationing.

Aug.—Interviewee commented, "Now I have a desire for a rug."

Sept.—No comment.

Oct.—Interviewee had purchased a rug shortly before the October interview. She stated, "Now I want a new refrigerator."

value of the desire answers exceeded the value of the resistance answers. Table 2 illustrates the responses obtained for an item which was not purchased—a car. Note the failure of the desire answers to increase relative to the resistance answers over time. In every case where the interviewee failed to purchase the item in question, the responses followed this general pattern.

The above cases were presented to illustrate the nature of the relationship between stated psychic tensions and purchasing behavior. Particular cases, however, give little evidence regarding the generality of this relationship. For this reason it was desired to work with an item which was purchased by all of the interviewees so as to permit for an averaging of results. Food was selected for this purpose.

Food. Since food is bought continually, it was analyzed in terms of whether or not changes in its rate of purchase might occur. This involved asking four questions, each to be answered in terms of the scale shown in Figure 1. Desire questions: (1) "How strong is your desire to increase your food items consumed (in quality or quantity) by 10 per cent from last month's amount?" (2) "How strong is your desire to avoid cutting your food items consumed (in quality or quantity) by

TABLE 2

PSYCHIC TENSION RELATIONSHIPS WHICH DID NOT
CULMINATE IN A PURCHASE—A CAR

(Data as obtained from Interviewee 69; comments made by him during the interviews are given below.)

	July	Aug.	Sept.	Oct.	Nov.	Dec.	Jan.
Desire for car	50	50	50	50	50-75	50	50
Resistance to expenditure for car	50	50	50	75	25-50	75	75

July—No comment.

Aug.—No comment.

Sept.—Interviewee commented that his present car had been repaired a few months earlier and he expected no further trouble for a couple of years, however it was beginning to look old. Consequently he might buy a car within a month or he might not do so for two years, but that he was watching ads conscientiously.

Oct.—Interviewee commented: "A new car looks more hopeless than ever." Earlier in the year he had agreed that his wife might buy a piece of furniture; she was now ready to purchase it and then he discovered that what she had in mind would cost around \$500.

Nov.—Interviewee commented that he was now thinking more strongly about a car because he had been inconvenienced a few days earlier by a breakdown with his present one.

Dec.—The interviewee seemed depressed. His wife had been taken to the hospital a few days before the interview and had had an expensive operation.

Jan.—No comment; no purchase.

10 per cent from last month's amount?" Resistance questions: (3) "How strong is your desire to reduce your food expenditures by 10 per cent from last month's amount?" (4) "How strong is your desire to avoid increasing your food expenditures by 10 per cent from last month's amount?"

To obtain the psychic tension ratings for food, answers to the above two desire questions were averaged and referred to as positive valences, and answers to the above resistance questions were averaged and referred to as negative valences. The arithmetic difference between these two magnitudes then was referred to as net valences (see Figure 2).

At this point a logical difficulty occurred: Are these psychic tensions sufficiently comparable interpersonally to permit for an averaging of the results from the interviewees concerned? To permit practical analysis, the following working postulate then was employed: If a homogeneous group of consumers all have an identical likelihood of making a particular purchase, a frequency distribution of their psychic tensions regarding the purchase in question will tend to form a normal proba-

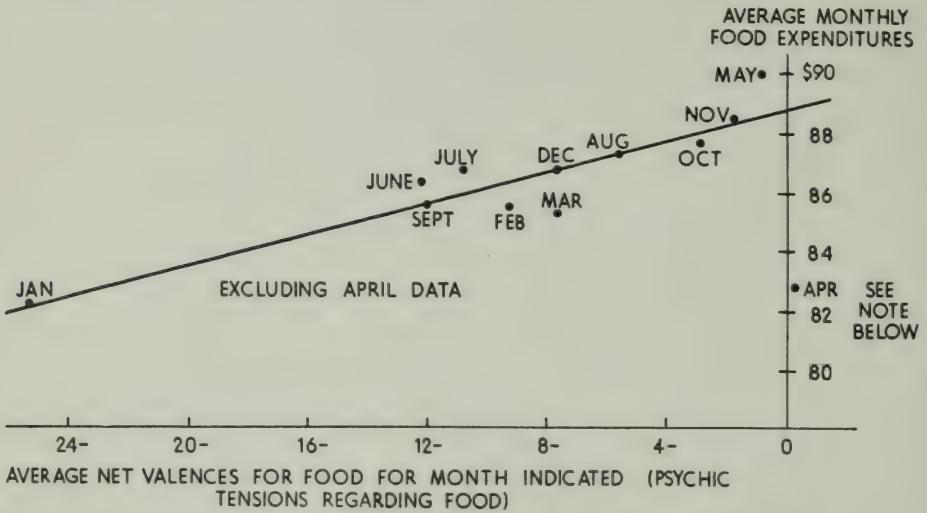


Fig. 2. Relationship between monthly averages of psychic tensions regarding food and average food expenditures for the same months. April data seems to be out of place because of inventory de-accumulation; many of the families stated that they closed out their lockers and finished eating their canned goods in April, so that their food consumption then was greater than their expenditures indicate. Excluding April data: $Y = 88.87 + .27X$ and $r = +.910$. Including April data: $Y = 87.51 + .17X$ and $r = +.550$.

bility curve. Although this working postulate has not been established, it does satisfy the pragmatic test of usefulness.

With the aid of the above working postulate the writer then averaged the psychic tensions (net valences, defined above) regarding food for all of the families by months. The corresponding averages of their food expenditures for the same months also were made. The results are shown in Figure 2. Note that the April relationship in Figure 2 seems to be out of place. That apparently was due to a general tendency on the part of many of these families to have reduced their food inventories during that month. A considerable number of the interviewees stated that their food expenditures were unusually low then because they had cleaned out their frozen food lockers in April. Most of them commented that they had eaten unusually well then. From the data shown in Figure 2, it appears that there is a rather close quanti-

tative relationship between the interviewees' psychic tensions for food and their corresponding food expenditures when allowance is made for inventory variation.

The analysis presented thus far has indicated the existence of a relationship between psychic tensions and purchases for both consumer durables and for food. The question then arose whether such a relationship might also hold for components within an expenditure category, for example, for particular food items. To test this, the interviewees were requested during the course of the study to maintain purchasing records for meat and eggs. Only a portion of them complied with this request, but the results from those who did keep such records indicate that a tension-expenditure relationship does exist (although less closely) for these items as well as for food as a whole. When cast in the form shown in Figure 2, the following relationships were obtained between the interviewees' psychic tensions and their corresponding monthly expenditures: for meat, $r = +.761$, and a regression line of $Y = 15.27 + .13X$; for eggs, $r = +.635$, and a regression line of $Y = 4.73 + .05X$.

The next step of the analysis was to discover factors accounting for these psychic tensions regarding food. The approach used was to correlate various co-existing and apparently relevant events with the desire and with the resistance components of this psychic conflict, each separately. By this means, it was found that the interviewees' desires were related to the disparity between their standard of living for food⁵ (measured as a matter of weighted attitudes—see Figure 3A, also following paragraph), and their plane of living for food⁶ (food expenditures deflated for price changes—see Figure 3B).

The term "weighted attitudes" refers to the interviewees' belief regarding how adversely they would be affected were their food purchases to be reduced by a given amount. This was used in Figure 3A as a means for approximating their standard of living for food. These weighted attitudes were obtained as follows. During the course of every interview each interviewee was asked, "Suppose you were to reduce your food items purchased (in quantity or quality) by 10 per cent from last month's amount, how great an effect do you believe

⁵ The "standard of living" for food is an economic term referring to that quantity and quality of food that the person concerned actively strives to obtain at the time in question. See J. S. Davis, "Standards and Content of Living," *American Economic Review*, Vol. 35 (March, 1945), pp. 1-15.

⁶ The term "plane of living" for food refers to that quantity and quality of food that the person actually obtains at the time in question. See P. Campbell, *The Consumer Interest* (New York: Harper, 1949); Davis, *op. cit.*

that this would have upon your family's: health? happiness? convenience? social prestige?" Each of these questions was answered by the interviewee pointing to a number on the thermometer scale shown in Figure 1. The numerical responses obtained then were treated as a

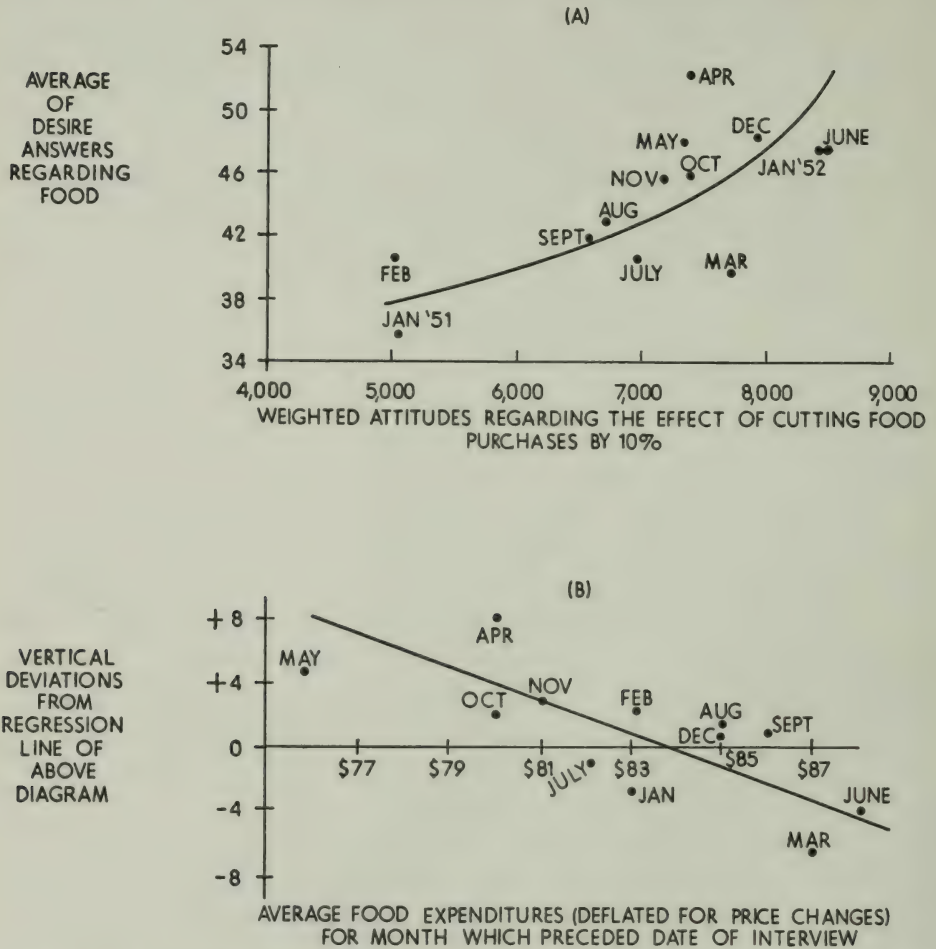


Fig. 3. Relationship between the magnitudes of the interviewees' stated desire for food, and (1) their corresponding weighted attitudes regarding the effect of a 10 per cent cut in food purchases and (2) their preceding month's food purchases. All data are averaged by months. Calculation is by the free-hand graphic multiple correlation procedure.

representation of the interviewee's attitudes regarding the effect which a 10 per cent cut in their food purchases would have upon her and her family. The next step was to ascertain the relative importance which the interviewees attached to health, happiness, convenience, and social prestige; to do this each was asked to indicate on the same scale (Figure 1) the importance to her of each of these factors. The average of

an interviewee's monthly ratings for each of these factors then was treated as the importance which she attached to it.⁷ Each interviewee's attitude answers then were multiplied by her corresponding "importance" answers, and the resulting products totaled so as to obtain her weighted attitudes regarding the effects of a 10 per cent cut in food purchases. These totals from all of the interviewees then were aver-

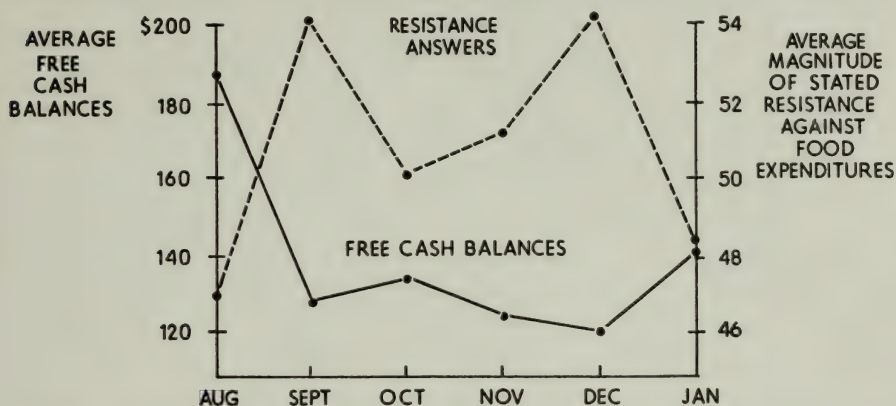


Fig. 4. Relationship between average free cash balances as of the beginning of each month and the average magnitude of stated resistance against food expenditures for the same month.

aged by months and used in Figure 3A as a numerical index of their standard of living for food.

The interviewees' resistance against making the expenditures which food purchases entailed were found to relate primarily to (1) the level of their free or uncommitted cash balances, and (2) changes in the level of food prices. To obtain the former, each interviewee was asked what he regarded as composing his free cash balance, and then to record the amount of this balance as of the beginning of each month. (Most of them stated that it included their cash in pocket plus their check book balance minus whatever was needed for payments on past commitments, such as insurance premiums, installment purchases, and rent.) Unfortunately, they were not asked to record these balances until August; but, for the period covered, the magnitude of the resistance answers regarding food was found to relate rather closely to the level of free cash balances (see Figure 4).

The best data available regarding food prices were the Bureau of

⁷ Unfortunately no attempt was made to determine whether this technique yields approximately the same results as the Allport-Vernon type scale might have given. See P. E. Vernon and G. W. Allport, "A Test for Personal Values," *Journal of Abnormal and Social Psychology*, Vol. 26 (October, 1931), pp. 231-48.

Labor Statistics' food price indexes for New Haven and Bridgeport, Connecticut. Figure 5 shows the relationship between changes in these indexes from one month to the next and changes in the magnitude of the interviewees' resistance answers regarding food for these same months. Note that the data in Figure 5 for October and June seem to be out of place. That is believed to be due to a discrepancy between changes in the Bridgeport and New Haven food prices and the

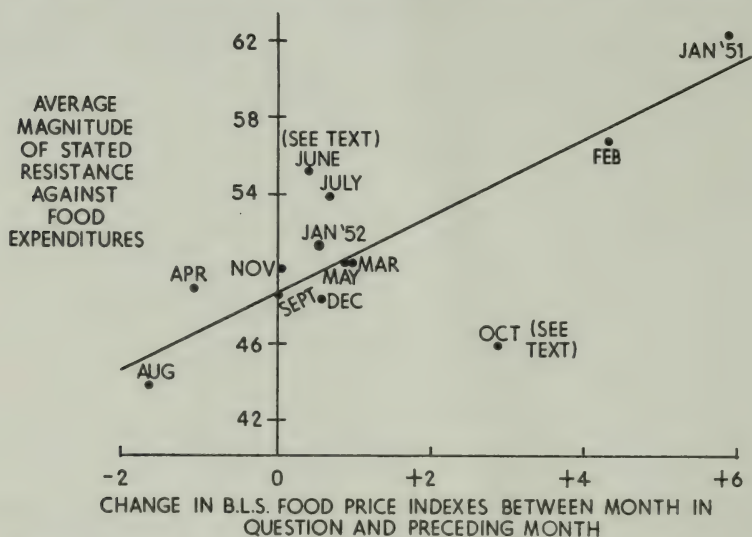


Fig. 5. Relationship between changes in B.L.S. food price indexes (average of New Haven and Bridgeport) and corresponding changes in the average magnitude of the interviewees' stated resistances against the expenditures which food purchases entail. Regression line is drawn free-hand.

Storrs food prices at those times. The reason for their being different then is that the Agricultural College of the University of Connecticut conducts meat-cutting classes between October and May of each year, and then sells this meat to faculty members at a price slightly below that prevailing in the open market. Thus, local food prices as a whole may actually have fallen from September to October even though the B.L.S. food price indexes rose between those months. For the same reason, the B.L.S. indexes may have risen by less than the local increase of aggregate food prices between May and June as the local interviewees shifted their meat purchases from the University to the regular private meat markets.

CONCLUSIONS

The results of this study indicate that there is a quantitative relationship between people's stated psychic tensions regarding the purchase of

particular items and their expenditures for those items. The further fact that these stated psychic tensions for food were found to relate to considerations (the interviewees' standard of living, their level of uncommitted cash balances, and to price changes) which economists have found by independent analysis to have an influence on consumer buying,⁸ indicates that the interviewees' stated psychic tensions tend to conform reasonably well with their actual psychic tensions—at least so far as food is concerned. This study therefore tends to validate Lewin's concept of conflicting psychic tensions and consumption behavior.

⁸ H. Schultz, *The Theory and Measurement of Demand* (Chicago: University of Chicago Press, 1938); U.S. Government, Board of Governors of the Federal Reserve System, *Federal Reserve Bulletin*, Vol. 32 (August, 1946), p. 847.

PART III

Nonprojective Survey
Techniques

INTRODUCTION

Psychological techniques are not the only means of ascertaining motivations. The last two parts of this volume are designed to present examples of other means of approach. In this part the emphasis is on methods that have been used to study motivations based on survey data. As will become evident from reading the selections in this part, a wide range of such methods are available. The selections incorporated herein are meant to convey the diversity of such methods both as to technique of approach and as to level of sophistication.

The selections in this part may be said to be of two types—those seeking to ascertain motivation directly and those seeking to ascertain motivation through the prior determination of attitudes, as through the use of rating or scaling methods. To some people, the inclusion of articles on measurement of attitudes in a book on determining motivation may seem rather surprising. Yet the fact is that there are many situations where it is more feasible, or more pertinent, to determine motivations by relating attitudes to certain indicators of purchase behavior than by probing for the “true” motivation by various interviewing techniques. In some cases, the correct identification of attitudes may be the *sine qua non* of the research problem, as is the case in evaluating the probable influence of alternative brand names in motivating the consumer toward or against the product.

For reasons such as these, the study of motivation and methods of determining attitudes are closely related, and a book on methods of identifying market motivation would hardly be complete without attention to attitude measurement. Some attention to this problem has been given in the previous part, in the article by Weschler and Bernberg discussing the measurement of attitudes by projective methods. More direct means of doing so are presented in three articles in this part of the book, the articles by George H. Brown and Seymour Banks presenting a more straightforward marketing approach and that by Louis Guttman describing one of the scaling methods for which he is so well known. These three articles by no means exhaust all the different approaches to attitude measurement, but combined with the Weschler-Bernberg article, they do provide some idea of the range of possibilities that are available.

These selections also bring out the fact that the more sophisticated

methods do not necessarily yield the most useful or the most reliable results. Techniques such as those described in the article by Allan Wilson can yield fully as significant results as the refined scaling method described by Louis Guttman. The important thing is to select the ones most appropriate in a particular case and use them correctly. Since each of the techniques described here is useful in some particular marketing situation, acquaintance with all of them is a worthwhile objective.

It is perhaps needless to note that the small selection of articles in this part cannot represent all the numerous nonprojective survey techniques that may be useful in the study of motivation. They should, however, provide some indication of the diversity of such methods.

21. A QUALITATIVE APPROACH*

The technique advanced in this article, the use of "informed opinion," is one that is likely to appeal to people who are neither psychologists nor statisticians, and may be a relief to those who may have begun to wonder if all motivation techniques were either psychological or statistical. This technique is a simple and informal one, though this does not preclude the need for advance thinking and planning. It is also a highly flexible technique, permitting modifications to be made at almost any stage of the field work, though this advantage is obtained not without careful revision in planning, extending even to modification of the original research objectives.

Note how the distinction between "how much" and "why" tends to become blurred in several of the illustrations. In a sense, this is inevitable for an understanding of why people make certain purchases often underlies quantitative market projections.

An important new development has taken place in marketing. It has happened quietly, and even some of those businessmen who have seen it have not become aware of its significance. Actually, qualitative market research, as the new development is called, is not complicated in principle, and it is not revolutionary. Indeed, the beginnings of the idea have been used by a number of businesses—including small ones—for several decades. It is, however, the recent elaboration and systema-

* Adapted from an article by Allan R. Wilson, sales and distribution consultant, "Qualitative Market Research," *Harvard Business Review*, Vol. 30 (January-February, 1952), pp. 75-86.

tizing of the techniques of qualitative market research which now will enable many managements to take a long step forward in finding markets for their products.

Management's ability to open new market frontiers will be more important than ever in a very short time. Government and business are currently working together in a large-scale expansion of productive capacity to meet emergency needs, but there has been no organized attempt to assure correspondingly expanded civilian markets when the military crisis is over. This is being left to private management, and qualitative market research is the most promising new tool with which to accomplish it. Here is an illustrative case from the last war:

A company developed a control device which was used in military aircraft. Its chief mechanical merits were that it allowed precise, vibration-proof adjustment of fuel supply with a minimum of friction and shock, controlled from a convenient position on the instrument panel.

The device was manufactured during the war under license. After V-J Day the manufacturing company's management felt that it had no postwar market, although the licensor was confident that wider applications could be found.

To reconcile the two differing viewpoints the manufacturer authorized an investigation, expecting it to confirm his belief that there was no postwar market for the device. To hold the expense of the investigation to a minimum, the survey was confined initially to the New York City area.

The market investigator took a decidedly qualitative, nonstatistical approach. He made up a tentative list of engineering consultants from a classified telephone directory. He called on these consultants and showed them the device, asking them where it might be used to solve any problems of their clients. They quickly appraised its characteristics and suggested the basic types of applications that could be made. These applications were developed in considerable technical detail.

As the interviews progressed, other consultants likely to be interested were added to the list, as were also companies where the device was thought by the consultants to have useful application. As a result, for example, a major railroad became interested. The device solved a difficult problem resulting from the effect of locomotive driving wheel vibration on the mechanical system of control cables leading from the wheels to the engineer's cab.

At the end of three weeks in New York City, several other possible applications became apparent, and a modest additional sum was authorized for a trip to Detroit for further investigation. At an automobile manufacturing company, the investigator found that the device could be installed as part of a dashboard emergency-brake unit, replacing a more cumbersome cable-control system. The new installation allowed a considerable unit-cost saving and had sales appeal in its greater convenience to women drivers.

Discovery of several other new uses for this ugly-duckling product required no more than 75 well-placed, intelligently directed exchanges of information with informed people.

During the war 50 people had been employed in the manufacture of the

device. At the war's end over 300 people were employed making the same product for which the management that authorized the investigation had felt there was no market!

Two points which stand out in this case are likely to apply to any story about qualitative market research. In the first place, the method used made it possible for the manufacturer to find customers whom he might never even have approached if he had relied on the traditional methods of the 1920's and 1930's. Secondly, the method made it possible for the manufacturer to contribute to a higher peace-time level of employment and business activity. If used by many firms instead of a few, the method is clearly of great potential significance to the economy as a whole.

In one sense, however, the story may be misleading. In retrospect, the techniques that made it a "success story" may seem simpler and more obvious than they really are. Actually, the kind of situation in which qualitative market research is particularly effective—a situation in which the market tides are muddy and changing—demands that the new method be used with skill and understanding. There is no simple formula.

Before examining the problem of using qualitative market research in detail, which is the principal purpose of this article, let us take a brief look at the status quo. How is market research carried on today?

MARKET RESEARCH TODAY

What we know and think about marketing today is the heritage of developments occurring principally during the past three decades. In the aftermath of World War I and throughout the 1920's, methods had to be developed to move the tremendous new outpouring of goods made possible by mass production. The depression of the 1930's generated a pressure for greater marketing efficiency that resulted in a range of developments from the supermarket and effective radio network advertising to better methods for gathering market information—especially the interview, questionnaire, and random sample.

The sales emphasis, it should be noted, was on nationally advertised and nationally distributed products of general use; and market research, based on assumptions of consumer market homogeneity that fitted in well with the growing structure of mass communications, was heavy with statistics of national wealth and trends of purchasing power in specific areas.

The change of emphasis to production, which took place during the 1940's, was significant in a less auspicious way. For most of this decade,

large manufacturers looked to government for decisions on what to make, how much, and at what price. There is a question as to how completely habits formed during the war and the immediate postwar boom were broken during 1949 and early 1950. For instance, how far beyond business transactions with the government did the idea of cost-plus selling featherbed itself? Indeed, there is at least circumstantial evidence indicating that the sellers' markets of the past decade have undermined serious interest in improving marketing methods.

Take, for example, the matter of membership in the American Marketing Association. The AMA is far and away the leading organization of its kind. Its standards are high, its journal is well regarded in professional circles; yet, according to the best figures I can find, only 263 of the leading 1,000 United States manufacturing concerns are represented by staff memberships in the AMA. Only one of the leading 100 banks is represented; only one of the 50 leading investment trusts; and only one of the 25 leading railroads. Even among the 50 largest retail store organizations only 20 companies are represented by membership. Moreover, the gap is not filled by economists (who in most firms where they are maintained devote a substantial portion of their duties to marketing problems). According to the National Industrial Conference Board, only 76 of the 1,000 leading manufacturing concerns maintain economists on their staffs.

It does not appear to be an unreasonable conclusion that the managements of many large-scale business operations have little assurance that they are adequately served by internal advice in either business economics or market research. Furthermore, the implications are very strong that the methods used by the great majority of businesses in gathering and evaluating their marketing information are either oversimplified or will not work for the more complicated problems of business administration.

Deficiencies in Present Methods. If we examine the common methods of gathering market information today, we are struck by a number of deficiencies. Current methods may be briefly summarized as follows:

The Accumulated Experience of Business Itself. This experience is most formally recorded in the internal accounting or statistical records of the business and in the written explanations of deviations from the standards set for performance. Much of the quantitative information gathered has as its basis the reporting of income for tax purposes.

Where these records are developed for the purpose of providing

genuine information, they are a useful guide in interpreting marketing changes. But too often they are regarded as merely an extension of a logical ladder of arithmetic ascending into the uniformities of a balance sheet and profit and loss statement. They become a static and tax-directed end, rather than a dynamic means to better marketing.

Another weakness lies in the fact that complex, unanticipated marketing situations often cannot be reduced to statistical terms until so much time has elapsed in accumulating the figures that the situation finally becomes of only historical interest.

External Statistical Guides Such as Trade and Government Reports. The usefulness of such guides has been considerable, particularly in pointing out uniformities by which individual businesses can compare their performance. A danger in the use of such statistics lies in the speed with which market conditions can change the situations which the statistics originally reflected. The market changes occurring during the second half of 1950 and the first half of 1951 created such crises for many companies—as in the case of a British company that stocked up heavily on raw materials for all its subsidiaries in anticipation of continued world-wide shortages on the basis of sterling-area statistics, only to find the picture changed by a very rapid recovery of Philippine production, not hitherto a major factor in the prewar supply.

Opinions of Department Heads, Employees (Particularly the Sales Force), and Supplier Salesmen. This method of getting information is still in common use, although to a considerable extent it has been supplemented by the use of the questionnaire and the application of statistical sales controls.

The use of such opinions as a basis for interpretation of marketing situations is widespread despite several obvious deficiencies. The judgment of each person is distorted by his own interests and by his narrow environment—distortions which are almost inevitable if the person's job is specialized—as well as by his limited ability to observe and to get reasonable, pertinent information.

While each of these opinions may reflect a fragment of a larger situation, the dimensions of which no single individual may be aware, there has been little effort to solve the problem of how to analyze and organize these fragmentary opinions into a cohesive and revealing picture.

The Periodic Use of Information Gathered by Means of the Questionnaire Used in Connection with the Random Sample. A large portion of marketing information is gathered by means of these techniques. They have widespread acceptance, and there is no doubting

their contribution to market research information. Nevertheless, the questionnaire has several important and decisive shortcomings. For one thing, even if it is very well prepared and the subject is very well selected, its results are not likely to be accurate during periods of public excitement such as elections and during wartime; its accuracy may be limited even by vacations and holidays. Again, even when the questionnaire method works relatively well, as often as not many questions will be badly misunderstood. In addition, varying local opinions, prejudices, and misconceptions will limit the accuracy of the information gathered.

But the major shortcoming of the questionnaire method is that it can never be used as a reliable basis for the most critical decisions of business management, that is, those caused by sudden, irregular changes in business and in the economy. The questionnaire is incapable of tracing out the unforeseen diversities that have entered into the new situations.

These shortcomings are most evident when questionnaires are distributed through the mail. Take the fact that the nonrespondents are quite likely to be those capable of giving the most useful information. Although such a difficulty can be avoided by sending out an interviewer with the questionnaire, it only leads to other shortcomings. The interviewing often has to be done by a student or by a part-time employee who is given minute instructions as to who shall be interviewed and under what conditions. The questions to be asked must be standardized. Under this type of semiskilled interviewing, there is little chance of locating unforeseen information which the questionnaire has not covered, yet which might bear on a really important business change. It would not turn up in the interview except accidentally; it would be only too easily misunderstood and misinterpreted by the interviewer; and the best to be hoped for would be that the interviewer brought back some indication that something was missed.

To be sure, the usual form of questionnaire or simple interview will do good service during the quiet, normal course of economic life. It will state accurately whether women prefer a gold-colored lipstick case to a green lipstick case, or whether women prefer candy or flowers as an anniversary present after 40. But, again, it will not provide the most needed and necessary advice as to any fundamental changes in business or business policies.

The inadequacy of these existing market research methods became particularly apparent in 1949 and pre-Korea 1950 when many businessmen became uneasy about piling up of inventories and shifts in

consumer demand. But if there was cause for uneasiness then, there is cause for alarm now. Some economists think it is quite possible that at the end of the present emergency expansion program the nation's productive potential will have doubled over that of 1945. Moreover, the new environment of a peace-war economy is bringing about new and unusual situations for marketing. On the one hand, military requirements are creating specialized markets for products ranging from light-weight oxygen tanks to waterproof radar units—markets governed by military standards of cost, utility, and consumption, and military purchasing procedures. On the other hand, there are the virtually infinite requirements of high-consumption civilian markets subject to the crosscurrents of fast-changing times.

So even though present methods of analyzing the market may have seemed adequate in the past, it is very doubtful that they will be able to cope with the future. The marketing problems we must face in the months ahead not only are bigger; they are also much more complicated.

THE USE OF "INFORMED OPINION"

It seems clear that a new method is now required to help solve the greatest problems affecting fundamental business decisions—those which harass the businessman under conditions of sudden or irregular change. Qualitative market research, which rests essentially on the gathering and evaluation of informed opinion, helps the businessman to see the change, meet the change, and utilize it to the best of his abilities. Indeed, this approach gives the greatest promise of lifting market research to its proper place as a tool of top management because it is best adapted to reflect the facets of a changing situation.

By gathering and evaluating informed opinion, we eliminate or minimize the need for the two major tools customarily employed in getting market information: the random sample and the questionnaire. We cannot use the random sample because informed opinion is not evenly distributed throughout the population, and because by its very nature it is a biased opinion. We can use the questionnaire only in broad form because we need maximum flexibility in recording the information if we are to allow the informed person to display every facet of his knowledge as he sees the situation. For example, the viewpoints of a sales manager faced with local marketing problems and of an advertising manager who is presenting a series of sales messages to a national audience may differ regarding the value of an advertising

campaign. Yet both may be right, and both may contribute equally to the final sales objective.

If we cannot exploit informed opinion to the necessary extent by means of the questionnaire, what alternative method is there for gathering, analyzing, and interpreting the information? The method I want to suggest is that of building up a series of case histories by means of interviews using a conversational sequence of questions and answers whereby the experiences of the interviewer and the interviewee are merged.

I am not thinking of so-called "depth interviewing" which, in the form used so extensively in the field of psychology, is concerned primarily with psychological attitudes, motivations, and reactions of individuals in relation to their own particular environments. In gathering informed opinion, we are concerned only indirectly with psychological aspects. The basic psychological motivations that make it possible to secure the informed opinion may be important (the businessman likes to talk about his problems, particularly to an understanding stranger), but they are not controlling.

Actually, beyond a superficial similarity in the method of gathering information by interview, there seems to be little in common between an interview conducted as psychological investigation and one conducted for the purpose of getting an informed opinion. Two entirely different points of view are involved: the environment of the psychological "depth interview" is predominantly emotional, while the environment of the informed opinion interview is predominantly rational and economic.

The process of using informed opinion in a marketing investigation can be covered in three steps: (1) preliminary fact-finding, (2) further detailed exploration of all the facets thus revealed, and (3) organization of the resulting information for management decision.

Preliminary Fact-Finding. Most executives have discovered that the business problem with which they start as often as not turns out to be something else before they solve it. This is particularly true of marketing problems, where the causes and effects are not observable in as tangible a form as are those on a factory floor or in a laboratory. Accordingly, the method of solving any marketing problem should not be considered airtight in advance.

The danger is, nevertheless, that the initial investigation will be limited by the wishes of executives who want the problem defined and handled according to their own particular ideas about its cause and

probable solution. Often, at the outset of a marketing investigation, there may be a wide range of experts on hand with definitions of the problem, each speaking with equal authority from vantage points in production, finance, product design, engineering, advertising, or sales. And each will have a long list of questions he wants answered to support his own particular view.

In order to accommodate the conflicting interests of those who wish their particular views verified in the market, a common procedure in building the questionnaire—and one which frequently goes astray—is to make a few market interviews that will tend to support or reject the ideas put forth by company executives concerning the direction that the investigation should take. To the extent that they are realistic, these interviews will tend to support the most accurate (or least inaccurate) definition of the problem. In order to avoid further controversy, interviews are then directed toward the very people who are likely to confirm the most clearly stated preliminary ideas about the nature of the problem.

In this procedure, the open questioning that might give valuable clues to the true nature of the problem often is cut off too soon. It is dropped before anything much more significant than descriptive or background material has been gathered. The result is that the subsequent investigation, frozen into a fixed list of questions, is long on the "how" and short on the "why" of the situation.

In contrast, when the marketing situation is genuinely explored through informed opinion, the definition of the problem is developed from interview to interview as the thread of the problem is found and traced through one set of experiences after another. When the persons most directly concerned with the problem are allowed to define it from their own individual points of view, each interview is strengthened by the widening net of information. Take, for instance, the following case concerning surgical instruments:

In a study of the buying influences affecting the purchase of surgical instruments, the investigators uncovered the reasons for buyers' specifying various types and qualities only after a detailed examination of a complex interrelationship.

In a hospital operating room there is an intricate combination of professional, technical, and housekeeping skills. The interviewer began with hospital purchasing agents, who indicated that surgical instruments of the highest quality generally were made of stainless steel. Operating-room supervisors confirmed that fact but added that stainless steel instruments were difficult to keep sharp. The surgeons indicated that this was true but that they preferred the workmanship of the instruments.

American surgical instrument manufacturers indicated that the best stainless steel instruments were imported, chiefly because of specialized difficulties in working that particular metal. They all said that other metals had more desirable working qualities and that other alloy steels had better characteristics for precise surgical work.

A recheck with surgeons indicated that they preferred stainless steel instruments principally because of a uniformly superior surface appearance over other steel instruments. Further conversations with operating-room supervisors revealed that stainless steel instruments did not corrode as did most other instruments under hard usage.

The reasons for corrosion brought the problem into focus. It was found that the majority of hospital instruments were sterilized in cleansing solutions requiring a close control over sterilizer temperatures. This routine had been developed originally when the sterilizing procedures in hospitals were performed directly by graduate operating-room nurses. As professional nurses upgraded their skills, and inherited work done formerly by interns, however, instrument sterilization was delegated to relatively unskilled hospital housekeeping personnel.

A wide range of sterilizer types, instrument cleaning solutions, and sterilizer control temperatures made this job mechanically—although not professionally—complicated. The result was that stainless steel surgical instruments were virtually the only instruments capable of withstanding the corrosive solutions created when sterilizers were operated by unskilled hands.

At the far end of this net of cause and effect, it was found that experimental research work in surgical teaching centers was being affected because of the prohibitive cost of manufacturing stainless steel instruments to special order.

This initial revelation of the interrelated reasons for preferences in instruments provided the basis for broader investigation in the same manner, which uncovered additional problems in the use of instruments. The latter phases traced out the influences of hospital accounting practices, of a trend toward group purchasing by hospitals, and of the hospital inspection standards set by national surgical organizations.

The entire investigation made use of no more than 40 interviews, although it required the informed and specialized viewpoints of the following types of individuals: hospital purchasing agents, hospital superintendents, group-insurance accountants, hospital group purchasing agents, operating-room supervisors, surgical chiefs of staff, hospital resident surgeons, and surgical instrument manufacturers.

Each of the 40 interviews averaged perhaps 50 questions and answers—a total of approximately 2,000 questions and answers. Even if 4,000 persons had been interviewed and asked a fixed list of 100 questions, determined in advance of the investigation by limited exploration—a total of 400,000 questions and answers—it is doubtful that the information gathered could have traced the problem even to a preliminary solution.

Further Detailed Exploration of All the Facets. The interview is essentially a case history, part of a series directed toward various aspects of a problem's solution. The total number of case histories required

depends not on the size of the sample—although this is a controversial point in market research—but on how sharply the people whose opinions are being sought are able to define their ideas about the problem.

There are good reasons why a very wide range of marketing problems can be solved from a framework of questions based on no more than 50 carefully chosen and properly developed interviews, in contrast to the partial solutions provided by hundreds or even thousands of superficial interviews. (Actually, 50 is not wholly an arbitrary figure; it has been confirmed as a rule of thumb by several investigators who have worked independently of each other on widely varied problems.¹) From the outset the informed opinion directs the investigation toward the people who are in a position to provide the problem's solution. The range of open questions and interrelated replies allows a practical definition of the limits of the investigation. The informed opinion keeps the initial investigation on the right track, pointing forward to other informed sources.

Note, also, that in this way advantage can be taken of unconventional sources. The real answer to a problem may be found in a seemingly totally unrelated quarter, such as would not be touched by a "tour of confirmation." Informed opinion often relates to the unknown or the exceptional as well as to the average situation. For example, the leading United States authority on the qualities of nuts and bolts is reputed to be a New England specialist in rebuilding machinery to meet unusual vibration problems; his most critical problems focus on nuts and bolts capable of standing up under severe stresses and strains. As another example, a roller bearing manufacturer is an expert on lubricants, because varying lubrication practices affect bearing maintenance and the long life of a wide range of types of machinery in which bearings are installed. A major objective of qualitative investigation is to track down the opinions of such men and give them and their counterparts adequate weight.

As interviews progress, certain facts and ideas begin to repeat themselves. After perhaps 20 interviews, ideas tend to group together in sufficiently clear form to indicate topics to be traced out in later interviews. This allows flexibility of questioning without getting far afield and permits later modification of the line of questioning as a particular trend becomes increasingly clear.

In this connection, think of the parlor game of "20 Questions." From any starting point a bright teen-age interrogator can pinpoint a

¹ See S. H. Britt, "The Strategy of Consumer Motivation," *Journal of Marketing*, Vol. 14 (April, 1950), p. 671.

hidden "fact" proposed by the other players. He finds the answer through a sequence of 20 "inspired" questions, each new one shaped in the light of the information previously revealed. In this game, a highly trained marketing executive armed with a fixed list of 20 questions made up in advance would be at quite a disadvantage compared to his teen-age competitor!

In the later stages of questioning in qualitative research, the various influences surrounding the problem will begin to take sharper form. Alternative courses of action may appear. For example, a new product developed for the consumer market may be discovered, during the course of investigation, to have stronger prospects in the industrial market.

At this point, particularly, there is a critical distinction between an open interview exploration of informed opinion and a limited exploration by means of a fixed list of questions. The fixed list may tell only whether a given thing can or cannot be done; the replies will give only a degree of approval or disapproval; and there is no way of discriminating between the quality of one opinion and another.

Organizing the Information for Management Decision. The actual organization of information develops naturally from the course of the interviews themselves, as the soundness of alternative courses of action is checked, opinion by opinion. Once the information has been organized for analysis, however, the interpretation is dependent as much on how the information is dealt with within the company as it is on the facts brought out in the market.

The analysis of a collection of informed opinions rests essentially on the nature of the ideas expressed. It is relatively easy to tabulate and analyze "yes" and "no" answers and settle them into proportions like colored beans in a jar. On the other hand, if opinions are being gathered among executives about what they consider to be the most important qualities of executive leadership, for example, the analysis becomes more complicated. If the discussions have really got down to cases, the breadth of the problem and the breadth of the ideas required for its solution will determine how much sorting and regrouping of ideas will have to be done. (Numerous coding systems have been developed, notably by the Department of Agriculture, to assist in the analysis of interviews in which some range of discussion is allowed within the categories of a fixed list of questions,² but information about purely qualitative analysis is meager.)

² H. E. Skott, "Attitude Research in the Department of Agriculture," *Public Opinion Quarterly*, Vol. 7 (Summer, 1943), pp. 280-92.

Perhaps as good a start as any toward analyzing a given collection of information is to look over a group of ten or so interviews dealing with the same product or function—food, drugs, machinery, advertising, machine design, accounting, or whatever it may be. What is said in repetition can be noted. This is done until the repetitions form a pattern. For example, in an investigation of the market for a paper-box machine, five shipping department executives in department stores emphasized that the range of box sizes produced by the machine was too broad for all but the very largest department stores. The repetition of the statement, backed up by logical reasons, was sufficient to eliminate the need for further investigation of department stores as a prime potential market.

As an illustration of interview analysis, here is the procedure that was followed in estimating the potential advertising revenue of a proposed publication in which a capital investment of over \$750,000 was at stake:

Opinions were secured from advertising managers, sales managers, advertising agency account executives, and media directors. The opinions covered a range of products from penicillin to locomotives. Each of these executives had something to say about why he considered that his company would or would not be a logical advertising prospect for the magazine. No interview guide, much less a fixed list of questions, could possibly have broadly covered all the seemingly divergent views.

After the discussions were completed, any idea or fact which seemed worth noting in any interview was entered on an individual 3" by 5" card. This card noted the job function, kind of company, and industry for each executive interviewed.

After all the facts and ideas were noted down, those which were similar were grouped together. These similar statements were then checked back against the discussions in which they had appeared originally. Out of a wide range of interviews and divergent opinions, several clearly defined points of view appeared. These ideas contained the basic issues of the problem. From them were derived standards for estimating the probable use of the publication by prospects rated in A, B, and C categories. For each group, individual estimates were made of the number of pages the prospects were likely to buy within a period of one, two, or three years.

All the reasons for using or not using the publication as an advertising medium added up to much more than votes of "yes" and "no." It was possible to evaluate the advertising volume of individual prospects with such accuracy that the over-all, first-year estimates of the magazine's advertising-page sales turned out to be 90 per cent of actual.

The interpretation of information does not necessarily depend on the method by which the facts have been gathered and analyzed.

Good judgment (if luck happens to be on its side) will overcome a bad method. Success, in the final analysis, depends on how the executive uses the information he gets.

Where a large-scale marketing change is concerned, some of the information about its nature can be prepared within the various departments affected. But the decisions for action concerning such changes must necessarily be made at the top. If the final management decision is left to an executive in a single department—sales, production, finance, or purchasing—a specialized interpretation may defeat the very purpose of the investigation, because the problem as finally demonstrated in all probability will cut across departmental lines.

Therefore, from the outset, the collection of information should be arranged to provide a guide for decision by top management, rather than to support a specialized point of view conditioned by day-to-day operations.

Typical Applications. Qualitative market research can be applied to many varying types of management problems, some of which have already been indicated. One valuable way in which the new method can be used is to indicate and analyze changing consumer trends earlier than other, conventional methods could. Here is an illustrative case:

In early 1949 there were approximately 450,000 television sets in the New York market area (compared to 2,435,000 sets by July 1951). Very little was known in early 1949 regarding television's impact on family living habits. Accordingly, a project was undertaken to build up this information.

The investigators felt that a thorough exploration of a limited number of families taken as a group would be more productive than a larger number of superficial interviews, as well as more economical than making separate records of individual activities. They counted on the fact that the common activities shared by adults, teen-agers, and small children as a group within the home—activities which would be affected directly by the television set—would also influence the more varied individual activities outside the home.

A group of 135 families was selected according to television-set sales in representative income areas. Open interviews were conducted with the family group in each home, exploring the changes they felt television had made in their way of living. Approximately 100 open questions and answers per interview, or a total of over 13,000 interrelated questions and replies, traced out the uniformities and differences in family response to television's impact.

The findings covered a wide range of home activity and showed the effect of television on such diverse factors as menu composition, recreation, home furnishing, and use of the family automobile. Problems of television programming also were brought into sharp relief.

In particular, the change in family recreational habits was so marked that it was decided to evaluate quantitatively the cumulative attendance by family

members at various recreational activities. These findings are summarized in Table 1.

Thus, valuable clues were found to entertainment attendance trends that were not confirmed by more conventional research methods until a year later, when television-set ownership in the New York area had increased to 1,100,000 sets. Numerous studies were then undertaken as a result of problems that

TABLE 1
EFFECT OF TELEVISION ON RECREATIONAL ACTIVITIES OF
METROPOLITAN NEW YORK FAMILIES

	Cumulative Attendance (Times per Month)		Per Cent Gain or Loss
	Before TV	Since TV	
Attended moving pictures	508	185	-64%
Attended theater (legitimate)	79	44	-32
Attended lodge meetings	85	70	-18
Attended night clubs	45	23	-49
Went bowling	86	76	-12
Attended boxing matches	62	34	-45
Attended wrestling matches	23	27	+22
Visited friends	408	289	-29
Entertained friends	427	773	+81

Source: *The Billboard*, June 18, 1949, p. 10. This survey was conducted by James E. Jump and Associates.

loomed over advertisers and entertainment executives. In almost every respect, the original 1949 study had clearly forecast problems that had to be taken up piecemeal at a later date.

In addition to making advance management planning possible, the qualitative method can be used to show how management can concentrate and channel its efforts more effectively. Here is the kind of thing that can be accomplished:

A national weekly news magazine selected a Midwestern city with widely diversified industries for a series of case histories of purchases made by business and industrial firms. The initial approach to each case was made through the company purchasing agent, who was asked for an account of the most difficult purchasing problem that had come to his attention during the past year.

It turned out that two major kinds of situation were: (a) those in which so many people were concerned in the purchase that the purchasing agent had continually to revise his specifications and prices, and (b) those in which he was not consulted until the decision to buy was virtually completed.

As these purchasing problems were traced out through groups of executives who had a part in the final buying decision, it was found that the method of processing purchases by purchasing departments was quite uniform for all companies. In contrast, the combinations of individuals influencing the final

buying decision did vary considerably from company to company and from purchase to purchase.

From a background of approximately 35 cases in the Midwestern city, plus the previous experience of an additional 25 cases covering another set of purchasing problems in another city, it was possible to demonstrate two types of business and industrial purchase, each with its own distinct pattern:

1. *Purchases Related Directly to the Production Line.* All purchases of raw materials, production machinery, and component parts, it was found, are regarded by buyers as bearing directly on the acceptance of the company's finished product in the market. Where such production purchases are concerned, the professional reputations of a wide range of production specialists are at stake in the tasks of design, development, and production. The interrelationship of skills in the production process and the interdependence of the scientific, engineering, and technical skills of the factory therefore lead to a minute examination and discussion of a multitude of details before the final decision to buy a raw material, a component part, or a production machine is made.

In bringing about such a final decision, the investigation revealed, personal salesmanship combined with considerable technical skill and knowledge is required in order to bring the supplying and buying companies into common agreement. This is particularly true because, as noted above, so many individuals may be involved—as many as 20 individuals in one buying company, for example, being directly concerned with a single change in a raw material specification.

2. *Purchases Not Related to the Production Line.* All purchases not fitting directly into a complex or planned manufacturing procedure, even when seemingly closely related (for example, the purchase of an electric motor driving a production machine or a lubricant oiling it), the investigation indicated, are not subject to the same intense personal or organizational scrutiny. Such purchases are not regarded as affecting the standards set for the product going through the manufacturing process. It is significant that nonproduction purchases of floor coverings, lighting fixtures, office equipment, materials-handling machinery, and the like are irregular or occasional, in contrast to the regular, repetitive pattern of production purchases. They are made after group recognition of their need, but the agreement that a need exists is relatively general and informal. Though purchases usually do require high executive approval, this appears in the main to be a formality held over from the close financial controls of the depression, and approval is rarely withheld.

As a matter of fact, the recognition of the need for a nonproduction purchase seems to develop almost spontaneously. Anyone from office boy to company president, for example, can recognize the need for a water cooler in a corner of the plant and set the wheels in motion for its purchase by putting a note in the suggestion box.

The distinction between the two types of purchases revealed by this case points toward two different combinations of selling and advertising in the promotional strategy of industrial-product manufacturers generally.

Because "production" purchases require continuous personal negotiation and readjustment of specifications between the seller and the manufacturing group, who are trying to agree among themselves on the final acceptable combination of men, materials, and machines, there must be a large element of selling effort; indeed such selling may be required almost continuously over a period of months. Conversely, advertising plays a subordinate, auxiliary role, except that assurance of quality and dependability is absolutely necessary and therefore advertising emphasizing the reputation of the supplier is a strong sales aid.

For "nonproduction" purchases, the role of advertising as a friend of the salesman appears to be primary (but in actual practice is greatly underestimated). While recognition of the need for such a purchase can originate with anyone within the buyers' organization, there is no guarantee that there will be even an informal agreement about the need, particularly one that can be relied on to last until a final decision has been made. And perhaps no more than one in ten purchases proposed for nonproduction-type products ever secures common agreement in the first place.

What can the salesman do? The cases showed that the salesman needs a system for getting clues indicating that someone in the organization has recognized the need for a new public address system or for new linoleum in the office lobby. But he cannot afford to spend his time ferreting out such needs by random prospect calls except at a prohibitive expense. The low-cost solution, it would appear, is through inquiries from advertising.

At the present time, however, there is an almost inexplicable lack of organized information as to what kinds of inquiries are likely to produce a sale and what kinds are not. Subsequent investigation has supported the conclusion that the salesman often regards the advertising inquiry not as a close friend (which it should be) but as an obstacle to improving his ratio of calls to completed sales.

Therefore management can take a step toward reducing the costs of selling for a wide range of advertised products by making a systematic appraisal of what characteristics of an advertising inquiry will most effectively lead to a sale. When this is done, it should be possible to establish a much better relationship between the costs of personal selling and advertising in the business and industrial marketing fields.

In any event, the evidence suggests that advertising in the business and industrial fields is most effective for products of occasional pur-

chase or for those bought over long or irregular intervals. Speaking more generally, the reduction of marketing costs by such controlled experiments as just described offers widespread opportunities both for profit and for improvement of marketing functions. At the same time, such activities offer the best means of further developing external standards of manufacturers' marketing costs.³

AREAS FOR FURTHER EXPLORATION

There are additional opportunities for applying qualitative market research methods, conditioned by both wartime and postwar developments. Several are suggested by changing marketing environments.

Getting Statements of Customer Intentions. Many customer lists have not been thoroughly examined in a decade. Salesmen turnover, expanding markets, and lack of necessity for determining the changing needs of customers have limited sales contacts within buyer organizations. In many accounts a long-term relationship has been reduced to a sole sales contact with the purchasing agent. Yet this relationship may have been determined originally by a combination of people who are now virtually unknown to the salesman.

Oversimplified buyer-seller contacts have also been created by jobbers and wholesalers who carry too many lines to go to the expense of ferreting out new influences on specific needs within buyer organizations. During 1951, lack of such information was responsible in part for situations where actual customer requirements were not accurately foreseen beyond the various log jams of speculative inventories.

Information about the intentions of customers cannot be secured by informal, irregular salesmen's reports, by quick executive trips "into the field," or by a routine listing of customer complaints. It can only be secured by an organization of up-to-the-minute information on what is really happening in customer organizations and, more important, why. This information must be gathered much more systematically and thoroughly than it has been in the past if it is to be effective.

In the advertising field, similar problems of appraising customer intentions face the sellers of publication space and broadcasting time. Media combinations are under new pressures not only because of changing local markets but also because of the localized market impact of television. Often there are also disparities between local sales objectives and national advertising objectives which require a recon-

³ See J. W. Culliton, "The Management Challenge of Marketing Costs," *Harvard Business Review*, Vol. 26 (January, 1948), pp. 88.

ciling of differences about the direction to be taken by merchandising and promotion strategy. These disparities cannot be identified and reconciled by mass-market measurements or statistical reporting devices. They require an appraisal of diverse viewpoints resulting from changing market environments.

Getting Sales Information for Investment Purposes. Corporate financing has been undertaken in the past with much less regard for the sales capacity of the companies financed than for the records of past earnings and the tangible assets on the balance sheet. A manufacturer's sales capacity may show no more tangible assets than office furniture and the personnel records of his sales staff. Yet the informed opinions of those who are familiar with the company's sales operation, from manufacturer to consumer, may be far more pertinent in assessing a company's future prospects than all the financial and production evidence of past success.

Under the provisions of the Securities and Exchange Act of 1933, information about a company's sales capacity and marketing competence can be demanded by the prudent investor. The fact is, however, that while various appraisals of sales capacity have been made for investment purposes, up to the present time the SEC has never required a company to file an investment prospectus in which its sales capacity has been evaluated.

The need for such information may be learned the hard way. During the past five years, two large stock issues were approved by the SEC and issued—one for a since-defunct automobile company and the other for a now equally defunct company organized to produce prefabricated housing—for companies in which there was little or no basic sales capacity or prospect of marketing success. It would seem to be only a matter of time before some company with equally dim marketing prospects will be challenged to prove that it has a selling mechanism comparable in adequacy to its production facilities.

At present one barrier to exhaustive marketing appraisals is the manner in which prime underwriters are designated by companies seeking financing. There is a very brief period—two or three weeks in many cases—between the time when prime underwriters are notified by the company that it is seeking funds and when the underwriting syndicate is formed. In this time, furthermore, hundreds of smaller securities dealers throughout the country must be notified that the securities are being offered for sale. Under these conditions the successful underwriter of the issue has little time to do more than work up an appraisal of the company's value from the standpoint of established

financial and accounting standards. If he can make any appraisal of sales capacity at all, it is necessarily a hurried one.

A tentative solution to this problem would be a rapid appraisal of informed opinion bearing on the company's sales standing derived from within the company's markets. Even though gathered under hurried conditions, such an appraisal would provide the prudent investor with at least a minimum of accurate information about the sales competence of the company in which he is considering making his investment—information to which he is entitled both by law and by good business sense.

New Product Development. There is a wide difference between the sums spent on manufacturing research and those spent for market research. There are reliable estimates indicating that marketing gets no more than 5 per cent of the total research funds spent by industry.

This unbalance in research investment is creating new pressures on marketing in many fields. Manufacturing research has at its disposal the world-wide developments of science in creating new combinations of materials and human skills. It is further supported by the virtually unlimited funds of government-financed military research programs. But this wealth of product development must then be funneled through a relatively inflexible marketing structure. The new products of a company's manufacturing research, for instance, may not fit into the sales structure for a multitude of reasons, ranging from executive whim to the price structure of competitors.

In addition, new products often are regarded as a burden by the sales organization unless it has specifically requested their development. The sales organization is held to cost-control standards that do not often allow large expenditures for investigating the prospective acceptance of new products by the market. Also, sales managers often have found that a new product developed because it would make a substantial contribution to manufacturing profit holds forth little promise of contributing to sales volume or sales profit. Therefore, in spite of the prospects for over-all company profit, many new products have not received the selling support originally expected of them.

Thus, the mortality in marketing new products is high because of the complex series of hurdles—often unseen at the beginning—that lie in their path. The specialized skills that go to make up the successful development and marketing of an established product are usually missing. Many of these skills are outside the company—in distributor, wholesaler, and dealer sales organizations. By continuously reappraising and re-evaluating informed opinion, however, we can effectively

bring the benefit of specialized skills to a new product going on the market.

22. MEASURING CONSUMER ATTITUDES TOWARD PRODUCTS*

Why marketing people should bother making attitude studies and the relevance of such studies to questions of market motivation is ably discussed in this article by George H. Brown. Three general types of applications of such studies to improving marketing techniques are brought out and supplemented with illustrative examples.

Three methods of measuring attitudes are also discussed and clearly explained in this article. These methods are by no means exhaustive, however, as will be apparent after reading some of the other selections in this part of the book.

The success of the continuous panel as a device for securing accurate sales data at the consumer level has led some students and practitioners of marketing research to question the wisdom of measuring consumer attitudes toward products. Sales measures obtained from research are easily understood and can be verified quickly by reference to other data. Measurements short of sales suffer the possibility that they may be unrelated to the fundamental objective of the marketing process, which is to maintain or increase the rate of sale for a particular product. What, then, are the unique gains that can be achieved through the measurement of consumer attitudes toward products?

As a first approach to this problem, let us consider the ways in which the measurement of attitudes can supplement the sales data available through a continuous panel. We know from experience and from panel data that the sale of individual brands within a product class are not equally distributed between each brand. Moreover, in many cases there is no tendency toward the sales becoming equal. This failure of brand sales to be equal reflects some fundamental factor at work, and one of the contributions of attitude research is to supplement panel data in explaining why a particular brand is doing poorly.

To show this point more clearly, we may list some of the more im-

* Adapted from an article by George H. Brown, formerly of the University of Chicago, now director of Consumer Research, Ford Division, Ford Motor Company, "Measuring Consumer Attitudes toward Products," *Journal of Marketing*, Vol. 14 (April, 1950), pp. 691-98.

portant factors that might cause a continuing inequality in brand sales in a particular market.

1. *The various brands of the product may not be equally available to the consumers.* It may well be that one brand of the commodity will be available at all retail outlets likely to carry such merchandise, whereas another brand may be available in only half or a third or even fewer of the available outlets.

2. *Consumer knowledge of the brands may not be equally distributed.* It is conceivable that a particular brand may be made available for short or long periods of time in a number of outlets in a given market, but that consumers may not be aware of the existence of the brand, solely as a result of its store display.

3. *Consumer opinions of the product attributes may differ markedly from brand to brand.* These product differences may be real or imaginary, but in any event the difference that exists in the mind of the consumer will reflect itself in the rate of consumer purchase.

4. *The price of the various brands may not be equal.* Other things being the same, experience has shown that the differing price explains differences in rates of purchase in the same market.

5. *A host of special factors may exist in any market situation.* One brand may be packaged in a different size package than competing brands. It may be possible to secure a premium as a result of purchasing one brand whereas a similar offer may not be available for competing brands, or if the offer is similar, the premium available may not be the same. One brand may have been introduced to the market earlier in point of time than another.

Consideration of the factors listed above will indicate that many of these possible sources of different brand popularities can be explained by data available through the continuing panel. Quite obviously, data on relative prices can be quickly and easily procured by either the store audit or by the consumer panel. The availability of the brand can be obtained directly in a store audit and at least partial data can be obtained through the consumer panel. Information on the package size, existence of premiums, special deals, and so forth can be easily obtained in panel surveys.

In spite of these substantial contributions, there still remains the problem of securing data on consumers' knowledge of the existence of brands and consumers' attitude towards brands. It is important for a manufacturer to know whether these factors are the cause of the lack of sale of his product, since the marketing action which he will take must be directed toward overcoming the obstacle to the sale. If the cause of the low sales is lack of knowledge of the brand among consumers, then the manufacturer must engage in advertising or some other form of communication in order to bring the sale of his brand to its proper place. If the difficulty is a low preference for the product,

then the manufacturer must investigate further to determine what attribute of the product or package is responsible for the opinion, and whether or not the attitude is based on a true evaluation of the product attribute.

If the consumers' beliefs are verified by objective tests, then the manufacturer must engage in product redesign or must accept the fact that the product attribute appeals only to a small segment of the market. If the attitude is based on a misunderstanding of the facts of the case, then advertising or other forms of sales promotion are indicated. Sales data do not give a full explanation of the cause of low sales. Some of the information gathered in connection with the sales data do throw light on this problem, but the possibilities of a direct measurement of consumers' attitude is a major factor in warranting continued experimentation with attitude research.

EXPLORATORY STUDIES OF CONSUMER ATTITUDES

To illustrate the point made above, a report will be made here on some exploratory studies undertaken at the University of Chicago with the support of General Mills. The purpose was to measure the knowledge level that exists in the market, as well as the attitude toward brands held by individuals who have had experience with the product. The techniques used in these measurements are well known to psychologists, and are similar to those used in marketing research by Cornelius DuBois.

The measurement of knowledge level was approached in two ways. First, individuals were asked to indicate the brand names which occurred to them when a particular product class was mentioned. The specific phraseology of the question was as follows: "When you think of ———, what brands do you think of?" In actual field tests, the degree of probing was carefully controlled in order to standardize the aid given to the respondent in answering such questions.

The second general approach to measuring the level of knowledge was the aided recall method where an individual is given a list of brand names and asked to state which brand or brands she is now using, which brand or brands have been used in the past, and which brands she has heard about. The tests included a fictitious brand to test the ability of the respondent to give the desired information. In the product classes used, however, consumers were well informed on their brands and did not claim to have used the fictitious brand. A modest percentage of the respondents claimed to have heard about the fictitious brand name, but the number was negligible.

Of the two methods, somewhat greater attention was given to the knowledge level based on the aided recall method than to the unaided recall. Effort was made to list all brands available in the market in the aided recall tests. By this device, answers were probably biased toward lesser known brands, but the alternative of deciding which brands to omit leads to acceptance of this limitation of the technique.

In measuring consumer attitude toward brands, the method of successive intervals—the familiar rating scale—was used as the basic technique. The method of paired comparisons and the method of rank order were also investigated, both of which are widely used in attitude measurement. The method of paired comparisons is generally looked upon as the basic measuring device, with the method of rank order and the method of successive intervals as approximations to the paired comparison data. The method of paired comparisons may be illustrated as follows: If consumer attitudes toward ten brands are desired, the brands would be presented two at a time to the consumer, with the question, "Which of these two brands do you like the better?" By getting an expression of opinion from each consumer for each of the possible pairs of brands, it is then possible to determine which brand holds the highest scale position in the market. In the case of ten brands there would be forty-five separate comparisons.

In the method of rank order, each person interviewed would be given all ten brands at one time and asked to sort them according to order of preference, placing the best liked brand in the number one position, and the least liked brand in the number ten position. No question other than the ordering between brands would be involved. These data can be converted to the paired comparison form by assuming that if Brand X rated one, and Brand Y is rated two, then on a paired comparison basis the individual would select Brand X rather than Brand Y.

In the method of successive intervals, the person is asked to take the ten brands and to indicate his attitude toward each brand by placing it in a category such as very satisfactory, satisfactory, indifferent, unsatisfactory, very unsatisfactory. Thus, the person not only ranks the brands, but indicates the intensity with which the preference is held. Moreover, by this method one or more brands may be placed in the same class, which does not occur under the paired comparison or ranked order methods.

The reason for selecting the method of successive intervals is that, of any of the three techniques, it enables an individual to handle the greatest number of observations. The interviewer fatigue that occurs

in any study arrives most quickly in the method of paired comparisons, then in the rank order test, and least in the successive intervals method. A number of experiments have been carried out, using all three methods, which indicate that the findings are approximately the same for each method. As a final factor, the method of successive intervals gives the most important opinion of the individual since it is assumed that the purchase will be concentrated on the item receiving the highest rating. As Thurstone pointed out in his classic article on the prediction of choice, it is necessary to go beyond the scale rating to the matter of distribution of first choices in order to make correct predictions.

To illustrate the application of these two measurements to particular situations, Table 1 shows the level of knowledge, preference rating, and rate of use of the leading brands of five different products in the Chicago market. These data were obtained, by the techniques described above, from two separate samples. The all-purpose flour, pie mix, and dry-soup data are from a sample of approximately 850 housewives selected at random from the Chicago metropolitan area during the month of May, 1949. The data for scouring cleanser and ice cream are from a sample of 500 housewives who were selected according to a probability sample design and interviewed during March, 1948. The data in the "now-using" column represent statements made by the persons interviewed, and may be used as an approximation of rates of sale.

Knowledge Levels and Preferences as Guides to Marketing Action. Table 1 shows the wide variation in knowledge level and preference rating that may exist in a market at a particular point of time. The all-purpose flour situation is particularly interesting. It shows that the knowledge level and preference ratings are approximately the same for the first three brands. The rate of use, however, shows a concentration in the first two brands, with a much smaller share of the market going to the third brand. These data alone indicate that some other factor, such as distribution or price, accounts for the comparatively low sale of the third brand. Moreover, the table suggests that neither of the leading brands can take action in this market to increase their sales substantially. Both products are well known, well liked, and purchased in approximately equal quantity. The most likely possibility for increasing sales would be to take some action causing a large increase in relative preference or to introduce a price variation either directly or through premiums.

The situation in regard to packaged pie mix is quite different from that of all-purpose flour. The knowledge level for the leading brand is well ahead of the next three brands, all of which are grouped around

TABLE 1
ATTITUDE TOWARD BRANDS IN SELECTED PRODUCT
CLASSES IN CHICAGO (SPRING, 1949)

<i>Brand</i>	<i>Per Cent of Housewives Who Know the Brand</i>	<i>Per Cent of Users Who Give Brand Top Rating</i>	<i>Per Cent of Housewives Now Using Brand</i>
<i>All-Purpose Flour</i>			
A	95.2	42.8	45.8
B	94.6	40.8	40.3
C	83.2	30.7	18.6
D	47.6	9.7	2.7
<i>Packaged Pie Mix</i>			
A	70.8	24.3	11.3
B	55.2	26.8	6.1
C	51.0	44.2	10.9
D	46.2	20.8	4.2
E	23.3	24.2	3.3
<i>Packaged Dry Soup</i>			
A	80.1	37.0	26.2
B	72.5	26.2	15.1
C	73.7	30.6	11.8
D	32.7	16.0	1.7
<i>Scouring Cleanser</i>			
A	95.1	45.7	41.2
B	62.8	44.9	16.3
C	43.6	29.4	11.0
D	36.0	46.7	10.0
E	65.8	24.6	6.1
<i>Ice Cream</i>			
A	33.8	67.5	21.4
B	51.8	67.6	17.8
C	26.8	50.0	8.6
D	16.8	53.6	3.7
E	37.8	26.5	3.2

a level of 50 per cent knowledge. The fifth brand is well below the second, third, and fourth brands and is known by approximately one third as many people as know the leading brand. In terms of preference rating, all the brands are about equally alike, with the exception of Brand C which stands well ahead of any other brand. The high preference for Brand C is shown by the fact that it has sales substantially in excess of Brand B or Brand D, which are known equally well. Brand A

does not enjoy the preference of B or C, but is nevertheless reportedly used by more people, reflecting the fact that more people are aware of the brand.

The marketing job to be done by the various manufacturers is fairly clear. All manufacturers except for Brand C have to investigate the quality of their product. If one brand such as C can command so strong a preference over all other brands, there must be some real or imaginary difference in the product quality. The manufacturer of Brand C can substantially improve his position in the market if he could increase the number of people who know about the product, and who have tried it. Advertising, sampling, counter displays, special promotions, and so forth, are clearly indicated since the evidence shows that sales will follow the use of this particular product. The manufacturer of Brand E can increase his share of the market by seeing to it that more people know of the product.

The possibilities of using data on consumer knowledge level and consumer attitude toward products in deciding what action should be taken by the manufacturer are clearly indicated by Table 1. The information should, of course, be supplemented by other types of information available in the market before planning a specific course of action. It should, moreover, be checked against data on distribution, price, package size, and any other types of information available from continuing panel studies. The important point is that attitude measurements can contribute to a general understanding of the market situation and can help in outlining the problem faced by the producer in holding a favorable market position or of overcoming an unfavorable position.

Attitudes as a Basis for Product Change. A second general application of the measurement of consumer attitude toward products is the guidance it gives in product development. As pointed out above, a low preference among consumers tends to be associated with a low market position. This fact, however, does not tell the manufacturer just what he should do to improve his product. Moreover, as indicated earlier in this article, it is quite possible that the attitudes of individuals may be in conflict with the facts, in which case the appropriate action would be advertising and sales promotion rather than changing the physical form of the product.

The approach to this aspect of the problem is to measure consumers' attitudes toward various attributes of the brands of the product class. Perhaps this can best be illustrated by a special analysis made on

scouring powders, in which individuals were not only asked to give their over-all rating, but were asked to give their opinion on each brand of scouring cleanser with reference to five product attributes.¹ The study showed that cleansing ability was the factor most highly related to over-all preference, with grittiness in use as the next most important factor. Package appearance, harshness on hands, odor, and price were much less important factors. This does not mean that package appearance or price and the other factors are unimportant, but that in the market situation being studied the products were considered more or less equal in regard to these product attributes, which were consequently not important in explaining the preferences that existed.

With this information in mind it was possible to make a more adequate analysis of the steps that might be taken by manufacturers faced with a low preference rating. Table 2 shows the average rating given

TABLE 2*
AVERAGE PREFERENCE SCORES ON PRODUCT ATTRIBUTES
OF SCOURING POWDERS, CHICAGO, SPRING, 1948
(500 interviews)

Product Attribute	Average Preference Rating		
	All Brands	Brand A	Brand C
Package appearance	.058	.037	.385
Cleansing ability	.066	.114	-.015
Grittiness in use	.058	-.424	.331
Harshness on hands	.133	-.162	.277
Odor	.329	.342	.331
Knowledge of price	1.795	2.171	1.823

* Data from Seymour Banks, *Some Applications of Psychological Measurement and Attitude to Market Research* (Ph.D. thesis, School of Business, University of Chicago).

to the various cleansers in the Chicago market. The outstanding points in this table are that Brand C rates very high on package appearance, both in relation to Brand A and to cleansers in general. On the other hand, Brand C rates somewhat lower on cleansing ability than Brand A, although it shows higher values for grittiness in use and harshness on hands. The negative values for Brand A indicate relatively low rating on the part of consumers in regard to grittiness and harshness, both with reference to Brand C and in comparison with all other cleansers.

In the case of Brand C, laboratory tests indicated that the cleansing

¹ This study is presented in the following article by Seymour Banks.—eds.

ability of the product was equal to or better than that of any other product on the market. The company producing this brand, however, had been emphasizing its package in current advertising, a factor which reflected in part in the high average rating on package appearance. We say "in part" because the company selected its package after exhaustive consumer tests and would have enjoyed a higher-than-average rating whether or not any advertising had been done. It appears, however, that the company might well have devoted its marketing energies toward correcting the attitude regarding the low cleansing ability of its product. Of course, better distribution and the possibilities of offering the product at a lower price would have contributed still further to its market position.

Measurement of Advertising Effects. The final use of measurement of consumer attitude toward products lies in the area of the measurement of advertising effect. Many producers have tried to use sales measures as guides to the effectiveness of their advertising activity. The influence of a single manufacturer's advertising, however, is frequently so small that the cost of securing the sales data necessary to measure the effect would be prohibitive, to say nothing of the costs of undertaking the appropriate analysis. By using attitude measurements to identify the marketing problem faced by the manufacturer, it becomes possible to define the advertising or marketing objective in such a way that the same basic technique can be used to determine whether or not the advertising or sales promotion effort is having the desired effect.

This type of approach to the measurement of advertising effect assumes that the attitude or preference changes before the rate of purchase of a particular brand changes. While this is a reasonable assumption, it has yet to be documented by careful research, and unless true it precludes the use of attitude measurements for this purpose.

COST OF ATTITUDE RESEARCH

Finally, attention should be called to the relatively low cost of attitude research. In the two studies reported here the field cost averaged slightly less than \$2.00 per completed interview, using an area probability sample. Experience has also shown that samples of 500 to 1,000 are adequate for the purposes described here and that data on seven product classes having a total of thirty major brands can be secured in one interview. This means that the field costs run from \$35 to \$70 per brand per market. It seems highly realistic to believe that on a co-

operative basis the full cost of providing these data should be less than \$200 per subscriber per study.²

23. WHY PEOPLE BUY PARTICULAR BRANDS*

This article serves to complement the preceding one by George H. Brown by discussing in detail one method of measuring consumer attitudes and relating the results to actual behavior.

That purchases can be explained in terms of brand preference would seem like a superficial explanation from a psychological point of view, but can nevertheless prove useful for marketing strategy, as is noted in this article. Note how the author also carries the analysis somewhat further through his use of discriminant analysis to measure the relative importance of various attributes of coffee and scouring powder in influencing both purchase of and preference for these products. Of course, this still leaves the author subject to the criticism of ignoring the erotic symbolism of these products . . .

The purpose of the research discussed below was the investigation of the relationships between preference and purchases by housewives of brands of seven classes of household products. Two bodies of data were collected and analyzed for correspondence: existing levels of preference for the brands of these products on the Chicago market, and purchases of these products by a panel of Chicago housewives during a three-week period.

COLLECTION OF DATA

The basic data were collected from a panel of 465 housewives in the city of Chicago during April-June, 1948. Each respondent was interviewed twice, the second interview coming three weeks after the first. During this three-week period, respondents reported purchases weekly by mail. The first interview was used to obtain information upon brands on hand, preference statements for brands, and a statement of purchase intentions for seven classes of household products. The

² In 1949 prices!—eds.

* Adapted from an article by Seymour Banks, Manager of Media Planning and Research, Leo Burnett Co., Inc., "The Relationships between Preference and Purchase of Brands," *Journal of Marketing*, Vol. 15 (October, 1950), pp. 145-57.

products used were scouring cleanser, coffee, ice cream, peanut butter, potato chips, mayonnaise and salad dressing, and catsup.¹

During the second interview, preference statements were collected again for brands of scouring cleanser and coffee. In addition, the respondents were asked to give preference scores on the product attributes of these brands. If comparison of mailed-in diaries and the original questionnaire showed that a housewife had failed to buy according to her stated intentions, she was asked why. Also, if on the first interview the respondent had expressed an intention of purchasing a lower-rated brand, she was asked the reason during the second interview.

A numerical rating scale for preferences was developed to avoid difficulties arising from differences of verbalization among respondents. It meant also that the respondent, not the interviewer, coded the respondent's answer. The scale used was read vertically and was portrayed in the shape of a thermometer. Numerals 0-8 were placed on the left side of the tube and phrases descriptive of the even numbered ratings were on the right. The highest rating was 8—very satisfactory; the lowest was 0—very unsatisfactory. Satisfactory was used as the basic word in the descriptions because this word conveyed the desired meaning—the ability of a product to produce utility for the housewife. The neutral point of the scale was placed at 2. The asymmetry was used because few pretest respondents had unfavorable impressions of brands that had been on the market for some time. In the rating procedure, the respondent held the preference scale and gave the numerical rating that best matched her feeling about a brand when that brand name was read to her by the interviewer. The respondents rated only the brands they had used.

The sample of housewives used in this study was drawn from the city of Chicago by an area sampling technique. Comparison of sample and population data showed the sample to be representative as checked by the following statistics: number of dwellings per block; percentage Negro; age of housewife; and years of school completed by housewife.

It was recognized early in the planning of this study that the procedure used to collect the data might seriously bias the results. The housewife who had bought something she had previously given a low rating might feel ashamed to admit it and so would report purchasing her most preferred brand to avoid "losing face." Doubt on this point

¹ The criteria used in selecting these products were: (a) wide use; (b) high rate of purchase; (c) equal availability to housewives; (d) absence of a brand completely dominating the market.

could be reduced substantially by obtaining objective measures of brand purchase by pantry audits, collection of labels or packages, and similar means. This would increase cost considerably, however, and would do nothing to keep a resolute respondent from cheating—she could always hide the bargain brand. To cope with this procedural bias, the following precautions were taken: preference and purchase data were obtained for many brands of seven different product classes; the interviewer was instructed to set a noncommittal tone with his respondent, "Just write down what you buy; a report of no purchase is a good report if that's what you did;" the diaries were to be mailed to the director of the study so the respondent would report her purchases to someone other than the person to whom she reported her preferences and purchase intentions; the diaries were mailed in weekly to reduce the amount of evidence available to the respondent to remind her of her statements or previous purchases. Still, the question of procedural bias is not completely settled; work remains to be done.²

DISCUSSION OF THE RELATIONS BETWEEN BRAND PREFERENCE AND PURCHASE

The basic data which were analyzed to determine the relations between preference and purchase were as follows, for each brand of each of the seven product classes: number of respondents who had used it and could remember it well enough to give it a preference rating, the number of those who had used it but could not remember much about it, and the number of those who had heard of the brand but had not used it—these numbers yield the total knowledge level for the brand; the proportion of 8's among raters; the brand's share of the panel's inventory of the corresponding product class at the first interview; the brand's share of the panel's purchase intentions for the product class; and the brand's share of the actual purchases of the product class. The distribution of actual purchases of brands by product classes was the datum plane for the analysis. All other data were compared in turn to the actual purchases.

The major interest of the study was in establishing the value of brand preference as a predictor of purchase. The basic hypothesis was that preference would show itself to be a good predictor of purchase.

² It is felt that the findings were not seriously affected by a procedural bias, if one remained. There was great variation in performance on purchase intentions by the respondents from product class to product class and by different individuals within the various product classes. Conversations with other researchers showed substantial agreement of the purchase data with reported sales of the brands of coffee and ice cream in Chicago at the time of study.

If this were upheld, the basic purpose of the work was accomplished. Two other possible predictors of purchase (past purchases as represented by brands on hand at the first interview, and future purchase intentions) were included in the data to serve as standards against which the predictive performance of preference might be compared. The information on the relative merits of these three as predictors of purchase was to be a by-product, not a main finding of the study.

The number of 8's among raters was chosen as the measure of preference for brands. This was chosen over several other possible indi-

TABLE 1
CORRELATION COEFFICIENTS OF PREDICTORS OF RELATIVE PURCHASE
OF BRANDS, BY PRODUCT CLASS

<i>Product Class</i> (1)	<i>No. of Brands</i> (2)	<i>Predictors of Relative Purchase</i>			
		<i>No. of Raters per Brand</i> (3)	<i>No. of 8's per Brand</i> (4)	<i>Brand's Share of Panel Inventory, 1st Interview</i> (5)	<i>Brand's Share of Purchase Intentions</i> (6)
Scouring cleanser	11	.828	.932	.985	.991
Coffee	20	.644	.670	.985	.898
Ice cream	15	.628	.750	.869	.912
Peanut butter	11	.945	.985	.997	.995
Potato chips	10	.853	.962	.987	.981
Mayonnaise and salad dressing	16	.859	.954	.995	.991
Carsup	13	.895	.974	.990	.992
Weighted average		.812	.918	.986	.977

cations of preference, proportion of 8's, a scale value, or an average rating. Almost all of the respondents intended buying only the brands rated 8. This meant lower ratings did not affect purchase decisions. Both scale values and average ratings include low ratings as well as high ratings. The proportion of 8's among raters is a measure of satisfaction with a brand; it could be used as a predictor of relative purchase of brands only when the numbers of raters for all brands were equal. This was not the case within the data. Nor can it be the case in the usual situation in which brands are on the market for varying amounts of time and with varying amounts of advertising and promotional effort.³

Table 1 summarizes the performance of the various predictors of purchase tested. The simplest predictor, the number of people rating each brand, was fairly good as a predictor of the relative purchase of brands.

³ Correlation of proportion of 8's per brand with purchases gave a weighted mean coefficient of .434 for the seven classes. This is much lower than any entry in Table 1.

But the number of 8's per brand, which takes into consideration the distribution of preferences for each brand as well as the number rating the brands, was significantly better in predicting relative purchase than the simple number of raters per brand. The number of raters per brand could equal the predictive performance of the number of 8's per brand only if the distributions of preference were the same for all brands. The basic data showed this not to be the case; there was great variation in the proportion of 8's from brand to brand in each of the product classes.

One might raise the question of the relation between the number of raters of a brand and the proportion of 8's given to that brand by its raters. Correlation of these two series gives a coefficient not significantly different from zero. Therefore, the better known brands were not more or less liked by their raters than the less well known brands.

The weighted average of column 4 of Table 1 is .918. This figure establishes the fact that there was a close relationship between preference (in terms of number of 8's per brand) and relative purchase of brands—84 per cent of the variation in relative purchase of brands is concomitant with variations in preference for those brands. The fluctuation of product-class coefficients of correlation in column 4 of Table 1 about their weighted mean is greater than might be expected from sampling fluctuation alone. The low r 's are for coffee and ice cream. Those products are characterized by large numbers of brands on the market and great variation in knowledge and use of those brands by panel members.

Two other items of data were tried as predictors of sales. These are last purchase made and future purchase intentions. Future purchase intentions were weighted by amount to be bought within the coming three weeks' period. This was the only weighted predictor used. The difference between the weighted averages of columns 5 and 6 is not significant. These last predictors of purchase were better than was number of 8's per brand. This is understandable. Preference measurements leave out important factors, such as price and accessibility of brands, which affect purchase behavior. Last purchase made and future purchase intentions reflect the impact of these other market factors, as well as preference, upon the housewife's actions.⁴

⁴ The basic purchase data used in the research discussed below were collected only for a three-week period. An obvious question of the findings is: over how long a period will a single statement of preference serve as an adequate predictor of purchase? To attempt an answer to this, preference data were collected from the families of my market research classes for brands of scouring cleanser and coffee, and weekly reports of purchase were made for fifteen weeks afterwards, from October, 1949, to January, 1950. The

Individual Performances. So far the discussion has been in terms of aggregates, and the question of how well the panel predicted the purchases of the entire group. The excellent performance obtained might have come about through every one doing exactly as forecast or by everyone doing the exact opposite, with the mistakes and errors offsetting one another. To investigate this, it was necessary to study individual performances, not group relationships.

There are two steps between preference and purchase. When one tabulates the purchase intentions of the members of the panel, a very close relation appears between preference, as measured by the highest rating used, and intention to purchase. Over the seven product classes, 96.4 per cent of the brands included in purchase intentions received a preference rating that was the highest or was tied with the highest the respondent used for that product class. The most common situation was for a housewife to give her highest preference rating to a single brand and name that brand in her purchase intentions; 65.8 per cent of the brands named in purchase situations were so rated. There was some variation in this from product class to class, with potato chips having the highest number of unitary preference-choice situations and catsup the least.

Relation of Purchase Intention to Actual Purchase. If there is a very close relation between the individual housewife's preference for brands and her purchase intentions, what is the relation between her purchase intentions and actual purchases? Considering all purchase intention statements, positive and negative, 52.4 per cent of the housewives' purchase intentions were carried out exactly: if they named a specific brand, they bought it; if they were going to buy from a group of named brands, not specifying which one, they bought from that group; if they said they were not going to buy anything during the

correlation coefficients of preference ratings in terms of number of 8's with combined purchase data for the fifteen-week period was .817 for the scouring cleanser and .841 for the coffee. These are at about the same levels as the early data which only covered three weeks. As in the earlier data, purchase intentions and past purchase data gave better predictive performances than number of 8's. The levels of predictive performance for these predictors were about the same height for each of the five three-week periods for both products except for that of past purchase with coffee. The distribution of brands of coffee found on the panel's shelves in October declined rather consistently in predictive ability throughout the study. It is not unexpected that in the fluctuating market that did obtain the distribution of brands on hand in October gave a poorer prediction of purchase in January than it did in November. The surprise is that the correlation was as good as it was: .79 in January compared to .96 in November.

These findings have test tube value only, because of the small size of the sample and its atypicalness. But the results do indicate the basic procedures described above may have ability to predict brand purchases over a much longer period of time than the three weeks of the major work discussed here.

time of the study, they reported no purchase.⁵ Apparently the housewives who did not intend to buy anything during the three weeks of the study were better prophets of their behavior than those who stated intentions to buy, 72.4 per cent of the former and 47.7 per cent of the latter type of purchase intention being sustained exactly. However, the performance, brandwise, on individual purchase intentions was really better than the last figure indicates, for there were many degrees of performance on brand purchase intentions (Table 2). Some of the

TABLE 2
PERFORMANCE ON BRAND PURCHASE INTENTIONS BY TYPE OF PURCHASE
INTENTION AND PRODUCT CLASS
(Data in percentage of respondents)

Product Class	Number	Type of Purchase Intention						
		Intended to Buy Some Brand(s)				Intended Not to Buy		
		Degree of Performance				Degree of Performance		
		Complete	Partial	Switched	Bought	Number	Complete	Failure
				Completely	Nothing			
Scouring cleanser	452	56.7	14.8	11.7	16.8	5	20.0	80.0
Coffee	442	60.2	22.9	11.5	5.4	18	55.6	44.4
Ice cream	337	32.3	19.6	29.7	18.4	101	61.4	38.6
Peanut butter	276	45.3	6.5	8.7	39.5	181	80.7	19.3
Potato chips	272	47.0	10.7	9.6	32.7	182	73.2	26.8
Mayonnaise and Salad dressing	409	43.0	16.1	16.7	24.2	33	68.0	32.0
Catsup	399	43.5	9.8	18.9	27.8	62	71.0	29.0
Weighted average		47.7	14.9	15.3	22.1		72.4	27.6

housewives carried out their purchase intentions partially either by buying only part of the brands mentioned or by buying their intended brands plus some others—14.9 per cent of those with brand purchase intentions fell into this class. Those who failed completely to fulfil their purchase intentions can be divided into two groups, those who failed to forecast correctly the time of purchase and those whose intentions and purchases do not agree at all, brandwise. It is only this last subgroup that might be considered as being really indifferent to brands or completely unable to predict their purchases. These people represent 15.3 per cent of those who gave brand purchase intentions and 12.4 per cent of the entire panel.

The impulsive nature of the purchase of ice cream is clearly il-

⁵ The following is a breakdown of the types of purchase intention found in the data: named specific brand(s), 71.0 per cent; named group without further specification, 10.1 per cent; and intended to buy nothing, 18.9 per cent.

lustrated; the smallest percentage of complete success on prediction of purchase and the highest amount of brand switching was reported for this product class. Coffee seems to be the product for which housewives did the best job of predicting brand purchases.

Preference Ratings of Brands Purchased. One can make a further study of the relationship between preference and purchase. This can be done by determining the preference ratings given to the brands actually purchased by the respondents. For the seven product classes over-all, the panel members bought four or five times as many 8-rated brands as they bought of lower-rated brands. Again ice cream showed up as a product in which there was less insistence upon preferred brands. This is only by contrast with the other product classes, for even in respect to ice cream the ratio of 8-rated brands purchased to lower-rated brands was three to one.

Purchases of previously nonrated brands bulked fairly high in the cases of ice cream and potato chips—the two chief examples of impulse goods in the study. It may be that the housewife was quite willing to buy unfamiliar brands of these products or that the interviewing procedure failed to elicit the full knowledge of the respondents. This latter is quite possible for respondents were not asked to give preference ratings to brands they had never used. If information on preferences of this group were obtained by the method DuBois uses,⁶ it is very likely that these figures of purchase of nonrated brands would be reduced.

Another way of determining the effect of preference upon purchase is to inquire into the motivation of those who intended purchases of less-preferred brands. This gives a look into those cases in which preference is not operative. Some of the respondents had stated intentions of buying brands of scouring cleanser and coffee they had given low ratings during the first interview. On the second interview, three weeks later, they were asked the reason for this intention. The reason for the intention was asked after the purchases, and not before, to avoid self-conscious action on the part of respondents.

Only a small number of housewives gave reasons for their previous intentions to buy lower-rated brands of scouring cleanser and coffee. The predominant reason was lower price on the lower-rated brand for both products. Other reasons, such as coupon offers, desire for change, and convenience got one or two mentions. The effect of price upon intention shows up quite strongly here. A question not answered by the

⁶ C. DuBois, *The Mind Is Bigger than the Pocketbook* (New York: Cornelius DuBois and Co., 1949), pamphlet.

preference measurement technique used is: How important is price in determining preference? These data gave evidence that in some cases price and preference are independent.

Failure to Make Intended Purchases. Preference affects purchase only through its effect upon purchase intentions. Failures of housewives to carry out their purchase intentions distorts the relation between preference and purchase. The second interview was used to ask those respondents whose mailed-in diaries showed failure to perform on their purchase intentions of brands of scouring cleanser and coffee why this had happened.

Leaving out those who had failed to buy because they had overestimated their needs, three main reasons were given for failure to buy brands of scouring cleanser and coffee as intended. They are price reduction on other brands, desire for change, and out of stock. The order is reversed for the two products. Of the three main reasons, price reductions on other brands was the most important reason for not buying scouring cleanser as intended and least important for coffee. "Out of stock" is the most important reason for buying other brands of coffee, and least important for scouring cleanser. Apparently the difference in housewives' minds between brands of cleanser is so small that small price changes will cause them to switch from one brand to another. The difference among coffees is greater; housewives switch brands only when the desired one is not available. The preference for a specified brand is probably not great enough for them to go to another store if the desired brand is not on hand at the first store. Coffee is still a convenience good.

The presence of desire for change is a limitation upon direct extrapolation of preference to predicted purchase. Because of its presence, the apparent and real choice situations of individuals are different. It is likely that satiation is a randomly occurring event and its effect will be cancelled out in a group purchase situation.

Analysis of Quantities Bought. As predictors of the number of units to be purchased, the panel fared poorly. Over-all, they bought 61.4 per cent more units of the seven product classes than they expected. The biggest error was for ice cream; the panels' members bought 2.6 times the predicted amount; the best job of quantitative prediction was for peanut butter, in which they exceeded their predictions by only 12.4 per cent. The errors by the respondents were not uniform in either direction or amount; 35.4 per cent of the predictions were correct, 38.8 per cent of the purchases were underestimated, and

22.3 per cent overestimated. The average error of underestimation was larger than the average error of overestimation.

THE EFFECT OF PREFERENCE FOR ATTRIBUTES OF BRANDS

The preceding discussion has related brand preference to purchase as if brands were indivisible objects. Brands are made up of many product attributes, however, some of which affect over-all preference and purchase strongly and others weakly or not at all. If a manufacturer wishes to increase sales by improving his product, he needs to know the most crucial attributes of his brand. The discussion which follows is based upon an exploratory effort to develop a technique for determining the relative importance of the various product attributes of a brand upon buyers' purchase decisions. The basic technique was multiple regression analysis of data for brands of scouring cleanser and coffee. Two types of studies were made, one on the relation between product attribute preference scores of brands and the over-all brand preference scores, and the other on the relation between brand attribute preference scores and brand purchase by the respondents. The major portion of the following discussion will be devoted to the latter type of study. Some comment upon the regression analysis of brand attribute ratings upon over-all brand preference scores can be found in an article by George H. Brown.⁷

In the first part of this article, the relation between brand preference scores and purchase was investigated by means of simple correlation. The obvious parallel for the study of the relationship between brand attribute preference scores and brand purchase is multiple regression, with quantity of purchase as the dependent variable and the attribute scores as the independent variables. The quantity of a brand purchased by a housewife, however, is affected by many things other than preference on product attributes, such as size of family, income, and shopping habits, for which no data were available. It was feared that omission of these variables would give regression functions that would do a poor job of fitting the data since most of the relevant variables were not included. To overcome this difficulty, it was decided to classify respondents into buyers and nonbuyers of brands, and find out how important these product attributes were in this separation. A technique known as linear discriminant function analysis has been developed to

⁷ G. H. Brown, "Measuring Consumer Attitudes toward Products," *Journal of Marketing*, Vol. 14 (April, 1950), pp. 691-98. Reproduced on pp. 268-277 of this volume.—eds.

sort objects into two classes on the basis of their measurements on several variables.⁸

The usual method of separating individuals into two groups on the basis of quantitative variables is to use a measurement on one variable at a time. The variable which gives the widest difference between means of the two groups is used for further work. This method is inefficient in that it cannot make use of several variables simultaneously. The linear discriminant function analysis allows the use of several variables at once to effect a separation into two classes and permits evaluation of each of the variables in this differentiation process.

The principal difference between a linear discriminant function and a linear regression function lies in the nature of the dependent variable. A linear regression function uses values of the dependent variable to determine a linear function that will estimate the values of the dependent variable. The discriminant function uses a two-way classification of the data to determine the linear function.⁹ Thus the dependent variable for regression is quantitative, the dependent variable for the discriminant is qualitative. The solution techniques of linear regression and linear discriminant functions are quite similar since they are both problems of variate analysis using linear functions. In both cases, equations are set up using variance and covariances of the independent variable scores and solved simultaneously.

The linear discriminant function is used to create an index number that determines which of two classifications an individual is to receive. The index number for each individual is calculated by summing the products of the scores of that individual on each of the independent variables multiplied by the coefficient determined for the strength of that variable in the function. The procedure for discrimination consists of finding some critical score such that index scores above this would result in the individual being classified with one group, and below this with the other. The weight given to each variable is a measure of its importance in this process of separation. These weights are what we are interested in for this study. They give the relative importance of each product attribute in separating the members of the

⁸ R. A. Fisher, "The Use of Multiple Measurements in Taxonomic Problems," *Annals of Eugenics*, Vol. 7 (September, 1936), pp. 179-88; D. Durand, *Risk Elements in Consumer Installment Financing: Technical Edition* (New York: National Bureau of Economic Research, 1941), Appendices A and B.

⁹ P. G. Hoel, *Introduction to Mathematical Statistics* (2nd ed.; New York: John Wiley and Sons, Inc., 1954), pp. 121-26.

panel into potential buyers and nonbuyers of brands of scouring cleanser and coffee.

Collection of Data. In the second interview, the respondent was asked to rate brands of scouring cleanser and coffee as she had done before. Then, ratings on attributes were obtained. All rated brands were scored on one attribute before going on to the next. This procedure was used to avoid a halo effect which would have obtained if all attributes of a brand were rated before going on to the next brand.¹⁰

The product attributes used for scouring cleanser and coffee were developed from free response questioning of a group of housewives. They were asked what interested them most when shopping for brands of scouring cleanser and coffee. The most frequently given responses were used as the product attributes of the study, as follows (these are not the order of frequency of mentions): Scouring cleanser—package appearance, cleansing ability, grittiness or scratchiness in use, harshness on hands, odor, and price; Coffee—package appearance, flavor, ability to make many cups per pound, and price.

In the regression analysis which follows, price was coded as 3 if the housewife knew the brand and 0 if she did not.¹¹

The attributes listed above are not unidimensional, nor are they independent. One housewife may give a brand of scouring cleanser 8 on grittiness because she likes the speedy way its coarse particles cut through dried food crusts. Another would give this brand 0 because she fears scratches on her porcelain sink from the coarse particles. The size of the partial regression coefficients were of more interest in this study than their sign. If flavor has a negative partial regression coefficient with brand preference or purchase for coffee, what should the manufacturer do, make the flavor stronger or weaker? The present data cannot tell. This work on fragmentation of brand preference is frankly exploratory. If it establishes certain attributes as critical for preference and purchase, it will have accomplished its purpose.

Before the preference data could be used for regression analysis, they had to be normalized. The original data were piled up around 8's, 6's and 0's. The data were coded by transforming the original values

¹⁰ Examination of the data leads one to the belief that a halo effect was avoided. The first order correlations within the scouring cleanser data between brand ratings and the product attribute ratings were tested for homogeneity. These coefficients are heterogeneous from product attribute to product attribute within brands and between brands. There is fairly high intercorrelation among the independent variables, however.

¹¹ Actually this procedure transforms a brand attribute into an attribute of the housewife. It was felt that she would be aware of the price if it was far out of line with other brands. The above process tests the ability of price to affect housewives' memory.

into values of 1, 0, -1, -2, -3. The zero corresponds to the original value of 7. The coded data are grouped within three standard deviations of their mean, whereas the raw data were very much skewed toward the low values.

Discussion of Results. Linear discriminant functions were computed, one relating the preference ratings on attributes of the major brands of scouring cleanser to purchase or lack of purchase of the eight leading brands of scouring cleanser, and the other doing the same with the eight major brands of coffee.¹² The coefficients of these functions were converted to units of the standard deviation of the corresponding variable—a process analogous to the computation of beta partial co-

TABLE 3
RELATIVE VALUES OF BETA PARTIAL REGRESSION COEFFICIENTS AND
DISCRIMINANT COEFFICIENTS OF SIX PRODUCT ATTRIBUTES ON
PREFERENCE AND PURCHASE OF BRANDS OF SCOURING
CLEANSER

(Pooled Data on 8 Leading Brands)

Case	Package Appearance	Cleansing Ability	Grittiness in Use	Harshness on Hands	Odor	Knowledge of Price
Preference	.10*	1.00*	.25*	.02	—	.10*
Purchase	.03	1.00*	.02	.27*	.14	.41*

* Original coefficients significantly different from 0 at the 1 per cent level of confidence.

Note: Entries converted from original coefficients, largest in each series given the value of 1.0 and others in relative terms.

efficients in multiple regression. In this form, the relative sizes of the coefficients come closest to stating the relative importance of the various product attributes in the discrimination process.¹³

Only three of the six product attributes of scouring cleanser were found to have partial discriminant coefficients significantly greater than zero. In order of relative size (and importance), they are cleansing ability, knowledge of price, and harshness on hands. Apparently, preference scores on package appearance, grittiness in use, and odor had no effect upon housewives' over-all preference ratings and purchase of brands of scouring cleanser. The bottom row of Table 3 shows the relative importance of the six product attributes for brand purchase. The top row of Table 3 summarizes the results of the regression analy-

¹² See P. O. Johnson, *Statistical Methods in Research* (New York: Prentice-Hall, Inc., 1949), pp. 347-53 for a discussion and demonstration of a solution procedure. He does not calculate the standard errors of the coefficients, but this can be done by means of the Fisher multipliers—Johnson, pp. 330-38.

¹³ Durand, *op. cit.*, p. 129.

sis of attribute preference scores of over-all brand preference ratings.¹⁴ Cleansing ability is the most important product attribute for both purchase and over-all brand preference rating; however, its dominance is different in the two cases. For preference, knowledge of price is only one-tenth as important as cleansing ability; but for purchase decisions, knowledge of price is 40 per cent as important as cleansing ability.

For coffee, only two of the four product attributes had partial discriminant coefficients significantly different from zero—flavor and knowledge of price. Table 4 summarizes for coffee the relative sizes of

TABLE 4
RELATIVE VALUES OF BETA PARTIAL REGRESSION COEFFICIENTS AND
DISCRIMINANT COEFFICIENTS OF FOUR PRODUCT ATTRIBUTES ON
PREFERENCE AND PURCHASE OF BRANDS OF COFFEE
(Pooled Data on 8 Leading Brands)

<i>Case</i>	<i>Package Appearance</i>	<i>Flavor</i>	<i>Ability to Make Many Cups per Pound</i>	<i>Knowledge of Price</i>
Preference	.03	1.00*	.10*	.05*
Purchase	.09	1.00*	.08	.43*

* Original coefficients significantly different from 0 at the 1 per cent level of confidence.

Note: Entries converted from original coefficients; largest in each series given the value of 1.0 and others in relative terms.

the attribute coefficients for purchase and contrasts them with the relative sizes of the partial regression coefficients for the attributes upon brand preference. As with cleansers, knowledge of price is several times more important in discriminating between buyers and nonbuyers of brands of coffee than it was in explaining variations in over-all brand preference scores. This finding is not surprising. It is comforting, however, to have intuition supported by evidence. This statistical confirmation of intuition underlines two statements: in considerable measure, preference and price are independent factors; and, preference and purchase situations are different.¹⁵

¹⁴ S. Banks, *Some Applications of Psychological Measurement and Attitude to Market Research* (Ph.D. thesis, University of Chicago), p. 113 (for the scouring cleanser data) and p. 115 (for coffee data).

¹⁵ The evidence for the last statement comes partially from Tables 3 and 4, which show the differing effects of the variables in the two situations, and partially from the multiple correlation coefficients of the two sets of equations. The multiple correlation coefficients of product attribute ratings with brand preference ratings are .836 for scouring cleanser and .882 for coffee; the multiple correlation coefficients for the discriminant analyses are .471 for the scouring cleanser and .537 for coffee. Thus a given set of variables does a much poorer job of explaining variation of brand purchase than it did in explaining variations of brand preference.

The findings on critical attributes cannot be exploited too far. Apparently, package appearance had no significant effect upon purchases of brands of scouring cleanser and coffee. This finding might have been an artifact of the research procedure. But even if it were not, a manufacturer should not feel free to package his coffee in torn, dirty bags. The results merely suggest that with present packages there is no apparent relation between preferences on this product attribute and purchase. This situation might have come about because all packages now in use are acceptable. Package appearances could come to have an effect upon purchase if some of the present packages were either improved or allowed to deteriorate.

CONCLUDING COMMENTS

On Methodology. All the information on brand and product attribute preferences used in this article was based upon unaided recall. The results might well have been different if the respondents were confronted with actual packages, used the cleansers, drank the various brands of coffee, and so on. Much research work needs to be done on the correct procedure for obtaining valid measurements of brand and product attribute preferences. Home tests with code-marked samples followed by preference tests and probing questions may be found to yield more valid data.

Possibly the product attributes used in this study need more refining than the method of measuring satisfaction upon them. They are chiefly verbal symbolizations of the product, and not single, unidimensional attributes. Attention must be given to the definition of attributes so that they are equally meaningful to respondents and to the production departments of manufacturers. Cleansing ability of scouring cleanser is based upon two factors, detergent action and abrasive action; the preference test should treat them separately and find which of these two factors is more important.

On Applications of the Research. The direct findings of this research may be summarized briefly as follows: there is a direct and close relationship between existing levels of preference for brands of seven classes of household products and their relative purchase by housewives. The first or highest choice is the most important factor in this preference-purchase relationship.¹⁶ By regression analysis, it was

¹⁶ The statement that only the first choice is important in choice is made by L. L. Thurstone, "The Prediction of Choice," *Psychometrika*, Vol. 10 (1945), pp. 237-53. The work in this paper supports this statement. If this finding is carried over into psychometric tests, it will simplify procedures now used to select the best or most preferred of a series of objects, such as illustrations or package designs.

possible to point out crucial attributes of scouring cleanser and coffee. It was found that different factors affect preference and purchase of brands. The factors which operate in both situations have different relative importance in the two situations.

The indirect findings of the study are the suggestions it offers for the application of the techniques and results to future research. Preference measurement can come to have a wider application than the prediction of the relative purchase of a series of competing alternatives. It can offer basic information as a guide to the formation of marketing policy. Market by market, the marketing manager can determine how well his product or brand is known, what people who have used it think of it, and the purchase intentions of users and nonusers. These findings will serve to point directly to basic marketing weaknesses. Remedies can be proposed for the weaknesses revealed by these diagnostic tests—advertising and sampling to increase knowledge level, product or package improvement to raise preferences for the product, and the use of price reductions or special promotional deals to increase purchase intentions if preference is high and purchase is low. If product improvement is suggested, over-all brand preference can be fragmented in terms of individual product attributes to determine the relative importance of these individual attributes in preference and purchase. These findings will permit the improvement in product to be directed more precisely than if it were guided by hunch or custom alone.¹⁷

Preference measurement can do more than serve as a means of discovering the basic marketing information mentioned above. It can also measure the response to marketing experiments designed to correct the deficiencies revealed by the first investigation. Thus, advertising campaigns can be evaluated in terms of their effect upon knowledge level and intentions to purchase. Product attribute changes can be evaluated in terms of product preference. Finally, change in knowledge level might be evaluated against change in preference as a means of increasing the purchase intentions for the brand in question.

A final consideration in this discussion on the use of preference measurement as a guide to marketing policy is the matter of costs to produce the desired change. If it becomes possible to evaluate the relative importance of knowledge level, product preference, merchandising factors, and price differentials in an equation, as it was possible to use product attribute preferences to predict over-all brand preferences, the problem is not ended when the relative importance of these factors

¹⁷ See Brown, *op. cit.*, for an extended comment on the use of preference measurement in diagnosis of a brand's market position.

is determined. For both the problem of market strategy and the problem of product improvement, the size of the coefficient in the regression equation is only part of the solution. It may be that some factors will respond more easily to a given amount of effort and money than others. Thus it may work out that a change of knowledge level is more important in affecting relative sales of a brand than changes in the preference level but the return in sales for a given investment is greater when a given sum of money is spent in increasing the preference for the brand through product improvement. For either change in factors outside the product or for changes of the product, the important consideration is the increase in sales per dollar expenditure. Preference measurement can serve to point out needed changes and to evaluate the performance of test solutions.

24. MEASURING THE TRUE STATE OF OPINION*

Finding out what people really think is an essential prerequisite to determining why people act as they do. Here is a direct-interview technique for ascertaining attitudes which has the unusual characteristic of being invariant of question wording. It also has the advantages of being easy to understand and of having been proved by considerable experience.

The technique has been developed and applied almost exclusively in public opinion work. Its applications to marketing are obvious, however, but have yet to be developed.

During the war a new approach to the problem of scaling attitudes and public opinion, called "scalogram" analysis, was developed by the writer to aid in the study of the morale and related aspects of the United States Army. This approach has wide ramifications not only for attitude and opinion research, but for many other fields like market research, mental testing, and elsewhere where it is desired to quantify qualitative data.¹

* Adapted from an article by Louis Guttman, formerly of Cornell University, now with the Israel Institute of Public Opinion, "The Cornell Technique for Scale and Intensity Analysis," *Educational and Psychological Measurement*, Vol. 7 (Summer, 1947), pp. 247-79.

¹ The basic concepts are available in L. Guttman, "A Basis for Scaling Qualitative Data," *American Sociological Review*, Vol. 9 (April, 1944), pp. 139-50. The work of the Research Branch of the Army Service Forces, done under the scientific leadership of Professor Samuel A. Stouffer, is described in several volumes published by the Princeton University Press. One of these volumes contains a rather comprehensive treatise on the theory and practice of scalogram analysis as carried out by the Research Branch.

The purpose of the present article is to describe another technique for scalogram analysis. Justification for the technique follows from the general theory and evidence published in the volumes on the Research Branch. We shall call it the "Cornell technique" for scalogram analysis to distinguish it from several alternative devices, since it was developed first for teaching purposes at Cornell. It is hoped that the reader may be able to master the technique from this present description.

The Scalogram Analysis Approach. The Cornell technique is a procedure for testing the hypothesis that a universe of qualitative data is a scale for a given population of people, using the scalogram approach. It may also be used to test the hypothesis that the data form a quasi-scale. Of the several techniques now available for scalogram analysis,² the one to be described here seems to be among the simplest and most convenient for general use. It requires no special equipment and involves only very simple clerical procedures which can readily be carried out by persons unskilled in statistics.

The various techniques just referred to all follow the same scalogram theory; they differ only in how the work is arranged. The initial steps are common to all. First, the universe of content to be studied is defined. In an attitude or opinion study, this means deciding on the general content of the questions to be asked. Second, the population of people is defined. In an attitude or opinion survey, this means that the class of people to be interviewed is delimited.

Next come two kinds of sampling problems. One kind is the ordinary problem of random sampling of people, and the other is the sampling of items. For these two sampling problems, it is helpful to distinguish between the pretest stage of a study and the final survey. Many fewer people can be used in a pretest than must be used in the final survey, but fewer items can be used in the final survey than must be used in the pretest.

In the pretest for a survey, about 100 persons will usually constitute an adequate sample of the population to test the hypothesis of scala-

² The first technique employed laborious least squares computations. See L. Guttman, "The Quantification of a Class of Attributes: A Theory and Method of Scale Construction," in P. Horst *et al.*, *The Prediction of Personal Adjustment* (New York: Social Science Research Council, No. 48, 1941), pp. 319-48. The standard procedure involves the use of scalogram boards especially invented for this purpose by the writer. A tabulation technique has been devised by another member of the Research Branch; see W. H. Goodenough, "A Technique for Scale Analysis," *Educational and Psychological Measurement*, Vol. 4 (Autumn, 1944), pp. 179-90. A brief statement of the Cornell technique as carried out on IBM equipment has already been noted in E. W. Noland, "Worker Attitudes and Industrial Absenteeism: A Statistical Appraisal," *American Sociological Review*, Vol. 10 (August, 1945), pp. 503-10.

bility. If the hypothesis is accepted, the items can be used in the final study to obtain reliable proportions at each scale rank.

The other sampling problem is of quite a different nature; it consists of sampling the universe of content. In an attitude or opinion survey, this is done by constructing some questions which contain the required general content. In a pretest, about a dozen questions usually can constitute an adequate sampling of the content. Since questions are constructed by the research workers, they do not fall into any standard random sampling scheme, and standard random sampling theory does not apply here. Instead, it is shown by the theory of scale analysis that almost any sample of about a dozen questions from the universe is adequate to test the hypothesis that the universe is scalable, provided the range of content desired is covered by the questions. If the hypothesis is accepted that the universe is scalable, then fewer questions can be used in the final study if fewer ranks are actually needed for the purposes of the final research.

Having defined the universe of content and the population of people, and having drawn a sample from each, the fifth step is to observe each person in the sample on each item or question in the sample. In an attitude or opinion survey where a questionnaire is used, this involves having the people indicate their answers to each question of the questionnaire.

The Hypothesis of Scalability. The problem now is to test the hypothesis, on the basis of the pretest sample data, that the entire universe of items forms a scale for the entire population of people. Let us review what this hypothesis implies.

The universe is said to be scalable for the population if it is possible to rank the people from high to low in such a fashion that from a person's rank alone we can reproduce his response to each of the items in a simple fashion.³ It is understood that a perfect scale is not to be expected in practice. Data have been considered sufficiently scalable if they are about 90 per cent reproducible, and if certain other conditions (to be explained later) are satisfied. For clarity, though, let us consider first a hypothetical perfect scale.

Suppose that a question is asked of a population concerning a certain political issue and that the responses are as follows:

Agree	60%
Undecided	10
Disagree	30
	<hr/> 100%

³ For a basic discussion of the theory of scales, see Guttman, "A Basis for Scaling Qualitative Data," *op. cit.*

If "Agree" means a more favorable opinion than "Undecided," if "Undecided" is more favorable than "Disagree," and if the universe is perfectly scalable, then the following must be true. The highest 60 per cent of the people must be those who said "Agree"; the next highest 10 per cent must be those who said "Undecided"; and the lowest 30 per cent must be those who "Disagree." If another question from this scalable universe is asked and the responses are 20 per cent "Yes" and 80 per cent "No," and if "Yes" means a more favorable attitude than "No," then the top 20 per cent of the people must be those who said "Yes" and the bottom 80 per cent must be those who said "No." From the rank of a person, we can now deduce what his response must be to each of these two questions. Any person in the top 20 per cent of the population must have said "Agree" to the first question and "Yes" to the second question. Any person lower than the top 20 per cent but not lower than the top 60 per cent said "Agree" to the first question and "No" to the second question. Any person below the top 60 per cent but not below the top 70 per cent said "Undecided" to the first question and "No" to the second, and the rest of the people, the bottom 30 per cent, said "Disagree" to the first question and "No" to the second.

The various techniques for scalogram analysis are devices to find the rank order which will best reproduce peoples' responses to each of the items in this fashion. If the universe were a perfect scale, all these techniques would involve little work and there would not be much to choose between them. It is the presence of imperfect reproducibility that raises the problem of technique.

The Cornell technique works by successive approximations. Usually just two approximations suffice to reject or accept the hypothesis of scalability. A first trial rank order for the people is established by a simple scoring scheme. For illustrative purposes, let us work out an actual case in detail. This illustration is not to be taken as a model of perfect research, but rather only to provide an example of the steps to be followed.

An Example of the Cornell Technique. It was desired to find out if the students in a class in race relations had a scalable attitude toward one of their textbooks, *A Nation of Nations*, by Louis Adamic. A questionnaire with seven questions was made out and administered to the class of 50 students. Both the number of questions and the number of students were smaller than those ordinarily used in a pretest; they were used here only because these smaller numbers permit displaying the full data.

The seven questions were as follows:

A Nation of Nations

Questions

1. *A Nation of Nations* does a good job of analyzing the ethnic groups in this country.
 ___Strongly agree(4) ___Agree(3) ___Undecided(2)
 ___Disagree(1) ___Strongly disagree(0)
2. On the whole, *A Nation of Nations* is not as good as most college textbooks.
 ___Strongly agree(0) ___Agree(1) ___Undecided(2)
 ___Disagree(3) ___Strongly disagree(4)
3. Adamic organizes and presents his material very well.
 ___Strongly agree(4) ___Agree(3) ___Undecided(2)
 ___Disagree(1) ___Strongly disagree(0)
4. As a sociological treatise, Adamic's book does not rate very high.
 ___Strongly agree(0) ___Agree(1) ___Undecided(2)
 ___Disagree(3) ___Strongly disagree(4)
5. Adamic does not discuss any one group in sufficient detail so that a student can obtain a real insight into problems of ethnic group relations in this country.
 ___Strongly agree(0) ___Agree(1) ___Undecided(2)
 ___Disagree(3) ___Strongly disagree(4)
6. By providing a panorama of various groups, *A Nation of Nations* lets the student get a good perspective on ethnic group relations in this country.
 ___Strongly agree(4) ___Agree(3) ___Undecided(2)
 ___Disagree(1) ___Strongly disagree(0)
7. *A Nation of Nations* is good enough to be kept as a textbook for this course.
 ___Strongly agree(4) ___Agree(3) ___Undecided(2)
 ___Disagree(1) ___Strongly disagree(0)

CONTENT SCALE ANALYSIS

We now describe, step by step, how the analysis of the responses is carried out by the Cornell technique.

1. Weights for the first trial are assigned to each category of each question, using the successive integers beginning with zero. In this example, since each set of answers has five categories, the weights range from 0 to 4. In each question, the higher weights are assigned to the categories judged to express a more favorable attitude. This judging of ranks of categories is not to be regarded as final; the consequent analysis will either verify the judging or determine how to revise it.

2. A total score is obtained for each person by adding up the weights of the categories he falls into. In our example, since the maximum weight for each person is four, and the total number of questions is seven, the total scores can range from 0 to 28.

3. The questionnaires are shuffled into rank order according to the total scores. In our example, we have arranged them from high to low.

4. A table is prepared like Table 1, with one column for each category of each question and one row for each person. Since each of our

questions has five categories, and since there are seven questions, we have 35 columns in our table. There are 50 students, so we have 50 rows. The first five columns are for the five categories of the first question, the second five columns for the five categories of the second question, and so on.

TABLE 1
FIRST TRIAL FOR CONTENT SCALE: *A Nation of Nations*

	1					2					3					4					5					6					7				
SCORE	4	3	2	1	0	4	3	2	1	0	4	3	2	1	0	4	3	2	1	0	4	3	2	1	0	4	3	2	1	0	4	3	2	1	0
28	X					X					X					X					X					X					X				
25	X					X					X					X					X					X					X				
25	X					X					X					X					X					X					X				
24	X					X					X					X					X					X					X				
23	X					X					X					X					X					X					X				
23	X					X					X					X					X					X					X				
22	X					X					X					X					X					X					X				
21	X					X					X					X					X					X					X				
21	X					X					X					X					X					X					X				
21	X					X					X					X					X					X					X				
21	X					X					X					X					X					X					X				
20	X					X					X					X					X					X					X				
20	X					X					X					X					X					X					X				
20	X					X					X					X					X					X					X				
20	X					X					X					X					X					X					X				
19	X					X					X					X					X					X					X				
19	X					X					X					X					X					X					X				
18	X					X					X					X					X					X					X				
18	X					X					X					X					X					X					X				
18	X					X					X					X					X					X					X				
18	X					X					X					X					X					X					X				
17	X					X					X					X					X					X					X				
17	X					X					X					X					X					X					X				
16	X					X					X					X					X					X					X				
16	X					X					X					X					X					X					X				
16	X					X					X					X					X					X					X				
16	X					X					X					X					X					X					X				
15	X					X					X					X					X					X					X				
15	X					X					X					X					X					X					X				
15	X					X					X					X					X					X					X				
14	X					X					X					X					X					X					X				
14	X					X					X					X					X					X					X				
13	X					X					X					X					X					X					X				
13	X					X					X					X					X					X					X				
12	X					X					X					X					X					X					X				
12	X					X					X					X					X					X					X				
11	X					X					X					X					X					X					X				
11	X					X					X					X					X					X					X				
10	X					X					X					X					X					X					X				
9	X					X					X					X					X					X					X				
8	X					X					X					X					X					X					X				
7	X					X					X					X					X					X					X				
7	X					X					X					X					X					X					X				
7	X					X					X					X					X					X					X				
6	X					X					X					X					X					X					X				
5	X					X					X					X					X					X					X				
5	X					X					X					X					X					X					X				
4	X					X					X					X					X					X					X				
FREQ.	9	27	2	12	0	8	24	0	13	5	10	25	8	7	0	3	7	16	14	10	3	14	5	21	7	9	21	7	12	1	11	19	5	11	4

5. The response of each person to each question is indicated on the table by placing an X in his row in the column for each category into which he falls. In our example, we have labeled the columns according to the questions and the weights of the categories. The first person is the one with the highest score, which is 28. He had checked the response weighted 4 in each of the questions, so he has seven X's in

his row, each under the respective columns for the categories with weight 4. There were two persons with a score of 25. The arrangement of people with the same score is arbitrary. Of the two persons in our example with a score of 25, the one placed first had a response of 4 to the first two questions, a response of 3 to the third question, of 4 to the fourth question, of 3 to the fifth and sixth questions, and of 4 to the seventh question. Similarly, the X's in Table 1 indicate the response of each of the remaining persons to each question. Every person answers every question,⁴ so that there are seven X's in each row. Table 1 gives a complete record of all the data obtained by the survey with respect to this area.

6. At the bottom of Table 1 are the frequencies of response for each category. Category 4 of question 1 had nine people in it, whereas category 3 of the same question had 27 people, and so forth. The sum of the frequencies of the five categories in each question is always the total number of people in the sample, which in this case is 50.

7. Now we come to the test for scalability. If the universe is a scale and if the order in which we have placed the people is the scale rank order, then the pattern of X's in Table 1 must be of a particularly simple kind. Let us consider the first question. If response 4 is higher than response 3, and if 3 is higher than 2, and if 2 is higher than 1 (response 0 happens to have no frequency in this case), then the nine people in category 4 should be the top nine people. Actually, six of them are the top six and the other three scatter farther down the column. Similarly, the twenty-seven people in category 3 should be below the first nine people and should go down to the thirty-sixth person ($36 = 9 + 27$). Again, this is not perfectly true for our data. A similar examination for the other items shows that there is a substantial error of reproducibility in their present form. The approximate number of errors need not be counted at this stage, since it is evidently more than 15 per cent of all the 350 responses ($350 = 7 \times 50$, the number of questions times the number of people) in Table 1.

8. It has seldom been found that an item with four or five categories will be sufficiently reproducible if the categories are regarded as distinct. One reason for this is the verbal habits of people. Some people may say "Strongly Agree" where others may say "Agree," whereas they have essentially the same position on the basic continuum but differ on an extraneous factor of verbal habits. By combining cate-

⁴ If people sometimes fail to respond to a question, then another category is added entitled "No Answer," which is weighted and treated like any other category for that question. In the present example, there were no "No Answers."

gories, minor extraneous variables of this kind can be minimized. By examining the overlapping of the X's within the columns of each question, it can be determined how best to combine the categories so as to minimize the error of reproducibility for the combinations. In question 2, for example, categories 4 and 3 seem to intertwine, so they are combined. Similarly, in the same question, categories 1 and 0 seem to intertwine, so they are combined. In question 4, on the other hand, we combine categories 3, 2, and 1, leaving categories 4 and 0 separate. The way to combine categories is determined for each question separately. The combinations decided upon for this example on the basis of Table 1 are given in Table 2.

TABLE 2
COMBINATIONS OF CATEGORIES

<i>Question</i>	<i>Combinations</i>
1	(4) (3) (2, 1, 0)
2	(4, 3) (2, 1, 0)
3	(4, 3, 2) (1, 0)
4	(4) (3, 2, 1) (0)
5	(4, 3, 2) (1, 0)
6	(4, 3) (2, 1, 0)
7	(4) (3) (2, 1, 0)

If it is desired to keep many scale types, then as little combination as possible should be done. If not many scale types are desired, however, the categories may be combined as far as one wishes even though this may not raise reproducibility. There is no harm in combining categories that could otherwise remain distinct with respect to scale error; all that is lost by such a combination is one scale type. On the other hand, categories may *require* combination in order to reduce error: they should be combined in the manner indicated by Table 1 and not arbitrarily.

9. A second trial rank order for the people can now be established on the basis of the combined categories. This is done by reassigning weights. Since the first question now has three categories (that is, three combinations), these are assigned the weights 0, 1, and 2. Question 2 now has two categories. These could be assigned the weights 0 and 1. In the present example the weights 0 and 2 are used instead, since keeping the range of weights relatively constant from item to item often helps to establish a better ranking for the people when there is error of reproducibility present.⁵

⁵ In a perfect scale, any set of weights, provided they have the proper rank order for the categories, will yield a perfect rank ordering for the people.

10. Each person is now given a new score, which represents his second trial rank order. This is done by rescoring his questionnaire according to the new weights. This rescoring is easily done from Table 1. Using a strip of paper which is as wide as the table, the new weights for the old categories can be written directly on the edge of the strip. Placing the strip across the row for a person, the weights are added according to where the X's lie. For our example, the strip would have for its first five columns the weights 2, 1, 0, 0, 0, weight 2 being placed in the column which was the old category 4, the weight 1 in the column which was the old category 3, and the 0's being in the old columns 2, 1, and 0 which are now combined. For question 2, the strip would have for the five columns the weights 2, 2, 0, 0, 0. Similarly, the new weights for the other questions can be written down to be used over the old columns of Table 1. The person who was formerly first on Table 1, with a score of 28, now has a score of $2 + 2 + 2 + 2 + 2 + 2 + 2 = 14$. The second person in Table 1 also gets a score of 14. The third person in Table 1 now gets a score of $2 + 2 + 2 + 1 + 2 + 2 + 2 = 13$; and so on for each person.

11. The people are now shifted into the rank order of their new scores, and Table 3 is prepared from the combined data just as Table 1 was prepared from the original data. Question 1 now has three columns, question 2 has two columns, and so on. The data of Table 1 are modified to fit Table 3 according to the combinations indicated in Table 2. The columns of Table 3 now refer to the combined categories, and the scores of Table 3 are the second trial scores just obtained in the preceding step.

12. The error of reproducibility in Table 3 seems much smaller than in Table 1, and we shall now count up the actual errors. This is done by establishing "cutting points" in the rank order of the people which separate them according to the categories in which they would fall if the scale were perfect. For question 1, which has three categories, we need two cutting points. The first seems to fall between the last person with score 12 and the first person with score 11. All people above this cutting point should be in category 2, and all people below should not be in category 2. Since there is one person in category 2 below this point, we have one error for category 2. A second cutting point is needed to separate category 1 from category 0; since these two categories overlap somewhat, its exact location is not essential since moving it slightly up or down will not change the amount of error. It should be placed so as to minimize the error, but this may be done in several adjacent ways. One way is to place the cutting point between

the second and third persons with score 4. Below this point we find three errors in category 1, and above this, we find five errors in category 0. The total number of errors in question 1 is $1 + 3 + 5 = 9$. Since there are 50 responses to question 1, this means 18 per cent

TABLE 3
SECOND TRIAL FOR CONTENT SCALE: *A Nation of Nations*

SCORE	1			2		3		4			5		6		7		
	2	1	0	2	0	2	0	2	1	0	2	0	2	0	2	1	0
14	x			x		x		x			x		x		x		
14	x			x		x		x			x		x		x		
13	x			x		x		x			x		x		x		
13	x			x		x		x			x		x		x		
13	x			x		x		x			x		x		x		
13	x			x		x		x			x		x		x		
12	x			x		x		x			x		x		x		
12	x			x		x		x			x		x		x		
11		x		x		x		x			x		x		x		
11		x		x		x		x			x		x		x		
11		x		x		x		x			x		x		x		
11			x	x		x		x			x		x		x		
11				x		x		x			x		x		x		
11				x		x		x			x		x		x		
11				x		x		x			x		x		x		
11				x		x		x			x		x		x		
11				x		x		x			x		x		x		
10				x		x		x			x		x		x		
10				x		x		x			x		x		x		
10				x		x		x			x		x		x		
9	x				x	x		x			x		x		x		
9		x		x		x		x			x		x		x		
9		x		x		x		x			x		x		x		
9		x		x		x		x			x		x		x		
9		x		x		x		x			x		x		x		
8		x		x		x		x			x		x		x		
7		x		x		x		x			x		x		x		
7		x		x		x		x			x		x		x		
7			x	x		x		x			x		x		x		
6			x	x		x		x			x		x		x		
6		x		x		x		x			x		x		x		
6		x		x		x		x			x		x		x		
6		x		x		x		x			x		x		x		
5			x	x		x		x			x		x		x		
4		x		x		x		x			x		x		x		
4		x		x		x		x			x		x		x		
4		x		x		x		x			x		x		x		
3		x		x		x		x			x		x		x		
3		x		x		x		x			x		x		x		
3		x		x		x		x			x		x		x		
2		x		x		x		x			x		x		x		
2		x		x		x		x			x		x		x		
2		x		x		x		x			x		x		x		
2		x		x		x		x			x		x		x		
1		x		x		x		x			x		x		x		
1		x		x		x		x			x		x		x		
1		x		x		x		x			x		x		x		
0		x		x		x		x			x		x		x		
FREQ.	9	27	14	32	18	43	7	3	37	10	22	28	30	20	11	19	20

error. This error could be reduced, of course, by combining the last two columns and leaving question 1 as a dichotomy. Then there would be only the one error in the first column. Such a further dichotomization need not be done if there is relatively little error in the other questions so that the error over all questions is not much more than 10 per cent.

Question 2 has two categories in the second trial, and the cutting point which will minimize the error is between the last two scores, which makes two errors in the first column and four errors in the second column of question 2. Similarly, question 3 has a cutting point between the last score 2 and the first score 1, leaving three errors in its second column. Question 4 gets two cutting points; questions 5 and 6, one cutting point; and question 7, two cutting points. The total number of errors in the whole of Table 3 is 40, which is 11 per cent of all the responses. We can, therefore, conclude in view of the fact that much of the error occurs in question 1 and could be eliminated by combining two categories in that question, that this area is scalable. From a person's rank order, we can reproduce his response to each question in terms of combined categories with 89 per cent accuracy (or better, if we combine the last two columns of question 1).

13. The per cent reproducibility alone is not sufficient to lead to the conclusion that the universe of content is scalable. The frequency of responses to each separate item must also be taken into account for a very simple reason. Reproducibility can be artificially high simply because one category in each item has a very high frequency. It can be proved that the reproducibility of an item can never be less than the largest frequency of its categories, regardless of whether the area is scalable or not. For example, question 3 in Table 3 has an extreme kind of distribution. Forty-three students are in one category, and seven in the other. Under no circumstances, then, could there be more than seven errors made on this item, regardless of whether or not a scale pattern existed. Or again, question 4 in Table 3 has thirty-seven cases in its modal category and thirteen cases in the other two categories. Under no circumstances, then, could item 4 have more than thirteen errors. Clearly, the more evenly the frequencies are distributed over the categories of a given item, the harder it is for reproducibility to be spuriously high. Questions 5 and 6 in Table 3 each have high reproducibility, each having five errors; these are not artificially high because question 5 has only twenty-eight cases in its more frequent category and question 6 has thirty cases for its modal frequency. The maximum possible error for question 5 is twenty-two, and for question 6 it is twenty. The scale pattern represents quite a substantial reduction from the maximum error. An empirical rule for judging the spuriousness of scale reproducibility has been adopted to be the following: no category should have more error in it than nonerror. Thus, the category with weight 2 in question 1 (Table 3) has eight nonerrors and one error; the category with weight 1 in this same question has twenty-four

nonerrors and three errors; category 0 has nine nonerrors and five errors. Thus question 1 fits this rule. Question 3 comes perilously near to not fitting the rule. While the first column of question 3 (in Table 3) has no error, the second column has three errors compared to four nonerrors. Similarly, the first column of question 4 has one error compared to two nonerrors. It is because evenly distributed questions like 5 and 6 have little error and because the errors in the other questions, like in 3 and 4, are not too widely displaced from where they ought to be, that we consider this area to be scalable.

In constructing a sample of items to be used in a test for scalability, at least some of the items should be constructed, if at all possible, to obtain a uniform distribution of frequencies. Such items afford a good test of scalability. Items with nonuniform frequencies are also needed, however, in order to get differentiated scale types, so both kinds of items must be used. The more categories that are retained in an item, the sharper is the test for scalability, because error—if it really should be there—has a better possibility to appear when there are more categories.

INTENSITY ANALYSIS

Separating "Favorable" from "Unfavorable" People. Since the expression of opinion about the textbook *A Nation of Nations* is sufficiently scalable, it is meaningful to say that one student likes the book better than another. There is a meaningful rank ordering of the students according to their opinion of the book. This ordering is expressed by the scale scores assigned in the second trial. A student with a higher score than another says the same or better things about the book (within scale error).

There is a further question that is of interest to the research worker. Given that the individuals can be ranked according to their degree of favorableness, is there a cutting point in this rank order such that we can say that all people to the right of the point are "favorable" and all people to the left are "unfavorable"? One person may be more favorable than another, yet both may be favorable. Obtaining just a rank order does not distinguish between being favorable and being unfavorable; it merely reflects being more favorable and less favorable and does not tell if a point is reached beyond which being less favorable actually means being "unfavorable."

An objective answer to this problem is provided by the use of the "intensity function." This function provides a solution to the traditional problem of question "bias." No matter how questions are worded or

"loaded," use of the intensity function will yield the same proportion of the group as favorable and unfavorable. The intensity function provides an invariant zero point for attitudes and opinions.

There are several techniques for obtaining intensity in a questionnaire, as is discussed in the volumes published on the work of the Research Branch. We shall discuss only two here, as carried out by the Cornell technique. These are very simple to perform. The first is the "fold-over" technique, and the second is the "two-part" technique. The fold-over technique is theoretically less justifiable than the two-part technique. It does have practical advantages, however, in some cases.

The Fold-over Technique. The fold-over technique consists simply of rescoring the content questions in order to obtain an intensity score. This is easily done for the form of question used to study opinions about *A Nation of Nations*. The following weights are assigned to the check list of answers: "Strongly agree" and "Strongly disagree" receive a weight of 2; "Agree" and "Disagree" receive a weight of 1; and "Undecided" receives a weight of 0.⁶ Thus, the apparently more intense responses receive higher weights, and the apparently less intense responses receive lower weights, regardless of whether the responses appear to be "favorable" or "unfavorable."

Weighting the responses in this way means that, in order to obtain an intensity score, we are in fact combining opposite ends of the check list, so that there are but three (combined) intensity categories per question. Intensity, as obtained in this fashion, is not in general scalable. Instead, it forms what is called a "quasi-scale." In a quasi-scale, there is no perfect relationship between a person's response to each question and his score on all the questions; instead, there is a gradient. The higher a person's score, the more likely he is to give a high response to each item, but there is not the high certainty that exists in the case of a scale. This can be seen in our example of Adamic's textbook. Arranging the data into a scalogram according to total intensity score, we obtain the configuration shown in Table 4. Each question now has three categories which represent the three intensity steps. There is a density gradient of responses. There are no clear-cut streaks in the category columns but, instead, gradually tapering densities that blend from one category into the next. Combining categories still will not yield a scalable pattern.

According to the basic theory of intensity analysis, intensity should

⁶ These weights can be written on a strip of paper to be put over Table 1 and added up there to obtain an intensity score for each person.

Plotting Intensity against Content. The empirical intensity function is obtained by plotting the intensity scores just obtained against the content scores obtained from the previous section from the second trial. The scattergram is shown in Table 5. The frequency in boldface in each column of Table 5 corresponds to the position of the median intensity for the respective columns. If the pure intrinsic intensity were being measured by our technique, there would be no scatter about these medians at all, but intensity would be a perfect U- or J-shaped function of the content scores. Despite the presence of error, however, the approximately shape of the true intensity function is clear from the shape of the curve along which the columnar medians lie. The curve

TABLE 5
SCATTERGRAM OF INTENSITY AND CONTENT:
A Nation of Nations

INTENSITY	CONTENT (SECOND TRIAL)							TOTAL
	0-2	3-5	6-8	9-10	11	12-13	14	
14							1	1
13								0
12	1							1
11				1		2	1	4
10					1	2		3
9	4	1	1		1	1		8
8	2	1	1	2				6
7	1	1	4	2	4	1		13
6		1	3	4	2			10
5		1						1
4		1			1			2
3		1						1
TOTAL	8	7	9	9	9	6	2	50

descends from the right, or the more favorable content scores, reaches its low point at the next to the last interval to the left (contents scores 3-5), and then rises again. The content scores 3-5, then, must be the approximate interval which contains the zero-point of the attitude. Students to the left of this interval can be said to have negative attitudes and students to the right can be said to have positive attitudes toward the textbook. Students in the 3-5 interval cannot be divided into positives and negatives without the aid of additional questions which will help to differentiate more precisely between their ranks.

On the basis of Table 5, we can conclude, then, that about 8 students did not like the textbook, 35 students did like the textbook, while 7 students were in between these. This division of the students into those with favorable and those with unfavorable attitudes does not depend upon the particular way we worded our questions. The same intensity curve, with the same proportion to the right and to the left of the zero-point, would have been obtained if we had used other questions or other wordings. provided only that these other questions were scalable

with the present questions. Proof of this invariant property of the intensity function is given in the volumes on the Research Branch's work.

Need for Larger Sample of People. This example must be regarded as a highly fortunate one, in one sense, for the purposes of this exposition. It is rare indeed to find as low error as we have in the intensity function so that the intensity curve and zero-point show clearly on the basis of our small sample of 50 cases. To perform an intensity analysis safely, when there is a substantial error present—which is the usual case—ordinarily from one to 3,000 cases are needed to obtain stable medians. To perform the scalogram analysis, it is also safer to use more than 50 cases. A hundred cases is a desirable minimum to use in the pretest, as well as a dozen or so items instead of seven as we have used in our illustrative example. If the pretest has established that the universe of items is scalable, the final study should be done on the usual number of cases used in opinion surveys if reliable results with respect to intensity are to be obtained. The hypothesis of scalability can be tested in a pretest on relatively few people because of its specialized character. Proportions of the population at any given rank or on one side of the zero-point, however, are subject to ordinary sampling error; larger samples of people must be used for reliable results.

Drawbacks to the Fold-over Technique. The fold-over technique for intensity has two theoretical drawbacks to it, as well as some practical ones. First, the intensity scores obtained thereby are not experimentally independent of the content scores, because exactly the same answers are used for both of the scores. This may give rise to some spuriousness in the relationship between the two. Second, it assumes that "Strongly agree" and "Strongly disagree" are approximately equal in intensity and opposite in direction, and similarly for "Agree" and "Disagree," while it is assumed that "Undecided" approximately straddles the zero-point. In fact, the occasional falsity of these assumptions is one contribution to error in the obtained intensity scores.

If the assumptions were true, life would be much easier for research workers. It would not be necessary to ask a series of questions in order to obtain a zero interval, because the "Undecided" category for any question would provide such an interval. But, unfortunately, it is clear that, in a series of questions on the same issue, the people who are "Undecided" on one question can all be "Agreed" on another question. It is just because we cannot interpret the bias of a question by looking at its content that such a technique like that of the intensity function is needed.

While the fold-over technique does have these two theoretical draw-

backs, it does seem to average out the errors involved in violating the above assumptions and to provide a proper U- or J-shaped curve in many cases.

A practical disadvantage to the fold-over technique has been found in the case of man-in-the-street interviews, where people avoid the "strongly" categories almost completely, so that not much differentiation in intensity can be obtained. In such a case, a two-part technique is necessary. An advantage of the fold-over over the two-part technique is that it takes less space and time in administering questionnaires. The two-part technique will be illustrated in the next example.

ANOTHER EXAMPLE OF CONTENT AND INTENSITY ANALYSIS

A Universe is Not Necessarily a Scale. A set of items constructed from a single universe of content is not necessarily scalable. The notion of universe of content and the notion of scalability are distinct. If a universe of content is not scalable, it can sometimes be broken down into subuniverses, some of which may be scalable separately. If a universe is not scalable for a given population of people, it is not meaningful to assign a single rank order to the people with respect to the total content. Indeed, if arbitrary scores were assigned to nonscalable data, intensity analysis should find that there was no U- or J-shaped intensity function and no invariant zero-point for dividing the population into positives and negatives.

An example of such a nonscalable case is the one next to be given. It will also illustrate the two-part intensity technique. The content for this second example concerns another textbook used in the same course as the first. The 50 students in the class were asked the following questions about *Black Metropolis* (by Drake and Cayton):

Black Metropolis

Questions

1. (a) On the whole, as textbooks go, how good do you think *Black Metropolis* is? (Check one answer)

<input type="checkbox"/> Very Good (5)	<input type="checkbox"/> Good (4)	<input type="checkbox"/> Fairly good (3)
<input type="checkbox"/> Passable (2)	<input type="checkbox"/> Not very good (1)	<input type="checkbox"/> Terrible (0)

 (b) How strongly do you feel about this? (Check one answer)

<input type="checkbox"/> Very strongly (3)	<input type="checkbox"/> Pretty strongly (2)
<input type="checkbox"/> Somewhat strongly (1)	<input type="checkbox"/> Not strongly at all (0)
2. (a) In your opinion, does *Black Metropolis* present a good sociological analysis of the Negro community in Chicago?

<input type="checkbox"/> An excellent analysis (5)	<input type="checkbox"/> A very good analysis (4)
<input type="checkbox"/> A pretty good analysis (3)	<input type="checkbox"/> It has only a few good points (2)
<input type="checkbox"/> Not a very good analysis (1)	<input type="checkbox"/> A pretty bad analysis (0)

 (b) How strongly do you feel about this?

<input type="checkbox"/> Very strongly (3)	<input type="checkbox"/> Pretty strongly (2)
<input type="checkbox"/> Somewhat strongly (1)	<input type="checkbox"/> Not strongly at all (0)

3. (a) To what extent does the book afford the student a real insight into the problems of race relations in Chicago?
☐ Not much at all (0) ☐ A somewhat limited insight (1)
☐ Fairly good insight (2) ☐ A good insight (3)
☐ A very good insight (4) ☐ An excellent insight (5)
 (b) How strongly do you feel about this?
☐ Very strongly (3) ☐ Pretty strongly (2)
☐ Somewhat strongly (1) ☐ Not strongly at all (0)
4. (a) In general, how well does the book organize and present its material?
☐ Very poorly (0) ☐ Not very well (1)
☐ Fairly well (2) ☐ Quite well (3) ☐ Very well (4)
 (b) How strongly do you feel about this?
☐ Very strongly (3) ☐ Pretty strongly (2)
☐ Somewhat strongly (1) ☐ Not strongly at all (0)
5. (a) Some parts of *Black Metropolis* emphasize statistical data and other parts quote personal interviews a great deal. Do you believe that the authors have succeeded in blending these data properly, or have they failed?
☐ Succeeded very well (4) ☐ Succeeded pretty well (3)
☐ Succeeded at least more than they have failed (2)
☐ Pretty much failed (1) ☐ Definitely failed (0)
 (b) How strongly do you feel about this?
☐ Very strongly (3) ☐ Pretty strongly (2)
☐ Somewhat strongly (1) ☐ Not strongly at all (0)
6. (a) Some students complain that the textbook often makes fuzzy statements, so that it is not clear what position it takes or what it is driving at. To what extent do you agree with this complaint?
☐ Completely agree (0) ☐ Agree for the most part (1)
☐ Undecided (2) ☐ Disagree (3) ☐ Completely disagree (4)
 (b) How strongly do you feel about this?
☐ Very strongly (3) ☐ Pretty strongly (2)
☐ Somewhat strongly (1) ☐ Not strongly at all (0)
7. (a) Do you think *Black Metropolis* is good enough to be kept as a textbook for this course?
☐ Definitely yes (4) ☐ Yes (3) ☐ Undecided (2)
☐ No (1) ☐ Definitely not (0)
 (b) How strongly do you feel about this?
☐ Very strongly (3) ☐ Pretty strongly (2)
☐ Somewhat strongly (1) ☐ Not strongly at all (0)

Each question is in two parts. The first part is to study content, and the second part is to study intensity. Notice that the number of categories in the content parts are not uniform from question to question. It is not essential for a scalogram analysis that there be any uniform format for the questions. In the same series of items, some can be trichotomies, some can have six categories, some can have two categories, and so forth. Nor is the wording of the categories of special importance. Short phrases, long phrases, and other variants can be used. Five and six categories were used in the present example because it was suspected in advance that the students would give apparently favorable answers to all questions put to them, so the apparently favorable responses were made more numerous in the check list of answers in order to help obtain differentiation in rankings.

The Cornell technique was used to analyze the content parts of the

seven questions on *Black Metropolis*. The first trial weights are those indicated with the questions, and the first trial scalogram is shown in Table 6. All of the items were found to have so much error that they required dichotomization. The combinations of categories used and the

TABLE 6
FIRST TRIAL FOR CONTENT: *Black Metropolis*

SCORE	1					2					3					4					5					6					7				
	5	4	3	2	1	0	5	4	3	2	1	0	5	4	3	2	1	0	4	3	2	1	0	4	3	2	1	0	4	3	2	1	0		
30	x						x						x						x					x											
29		x					x						x						x					x											
28	x						x						x						x					x											
28	x							x					x						x					x											
28	x								x				x						x					x											
27		x								x			x						x					x											
27	x							x					x						x					x											
27	x								x				x						x					x											
27	x									x			x						x					x											
26		x									x		x						x					x											
26	x												x						x					x											
26	x												x						x					x											
26	x												x						x					x											
26	x												x						x					x											
26	x												x						x					x											
25	x												x						x					x											
25	x												x						x					x											
25	x												x						x					x											
25	x												x						x					x											
24	x												x						x					x											
24		x											x						x					x											
24		x											x						x					x											
24		x											x						x					x											
24			x										x						x					x											
24				x									x						x					x											
24					x								x						x					x											
23			x										x						x					x											
23				x									x						x					x											
23					x								x						x					x											
23						x							x						x					x											
22							x						x						x					x											
22								x					x						x					x											
22									x				x						x					x											
22										x			x						x					x											
21											x		x						x					x											
21												x							x					x											
21													x						x					x											
21													x						x					x											
20													x						x					x											
20													x						x					x											
19													x						x					x											
19													x						x					x											
19													x						x					x											
18													x						x					x											
16													x						x					x											
16													x						x					x											
14													x						x					x											
14													x						x					x											
13													x						x					x											
FREQ.	18	19	10	2	1	0	12	26	12	0	0	0	12	22	10	4	2	0	14	21	9	5	1	16	26	6	2	0	2	19	12	15	2	20	

results of the second trial are shown in Table 7. There is still too much error in Table 7. Several of the questions have more error than non-error. We therefore judge the total content not to be scalable, since no further trials can be made when all items are dichotomous.

Therefore, we cannot speak of degrees of "favorableness" of opinion about *Black Metropolis* for this class of students. We cannot say that one student likes the book better than another student. He may like it better in one of the aspects and not better in another. There is ap-

it anyhow. Each part (b) of the 7 questions on *Black Metropolis* was weighted according to the weights indicated in the list of questions above, and trial intensity scores were obtained thereby. Intensity again seems to be a quasi-scale. Obtaining a quasi-scale, however, has no

TABLE 8
SCALOGRAM OF INTENSITY: *Black Metropolis*

SCORE	1				2				3				4				5				6				7			
	3	2	1	0	3	2	1	0	3	2	1	0	3	2	1	0	3	2	1	0	3	2	1	0	3	2	1	0
21	x				x				x				x				x				x				x			
20		x			x				x				x				x				x				x			
20			x		x				x				x				x				x				x			
20	x				x				x				x				x				x				x			
19	x				x				x				x				x				x				x			
18	x				x				x				x				x				x				x			
18		x				x			x					x			x					x			x			
18	x				x					x			x				x				x				x			
17		x			x				x				x				x				x				x			
17	x				x				x				x				x				x				x			
17		x			x				x				x				x				x				x			
17	x				x				x				x				x				x				x			
17		x			x				x				x				x				x				x			
16		x			x				x				x				x				x				x			
16			x			x			x					x			x					x			x			
16	x				x				x				x				x				x				x			
16		x				x			x					x			x					x			x			
16	x				x				x				x				x				x				x			
16		x				x			x				x				x				x				x			
15			x				x		x				x				x				x				x			
15				x					x				x				x				x				x			
15					x				x				x				x				x				x			
15						x			x				x				x				x				x			
15							x		x				x				x				x				x			
15								x	x				x				x				x				x			
14								x	x				x				x				x				x			
14									x				x				x				x				x			
14									x				x				x				x				x			
14									x				x				x				x				x			
13									x				x				x				x				x			
13									x				x				x				x				x			
13									x				x				x				x				x			
13									x				x				x				x				x			
13									x				x				x				x				x			
13									x				x				x				x				x			
12									x				x				x				x				x			
12									x				x				x				x				x			
12									x				x				x				x				x			
11									x				x				x				x				x			
11									x				x				x				x				x			
11									x				x				x				x				x			
11									x				x				x				x				x			
11									x				x				x				x				x			
10									x				x				x				x				x			
10									x				x				x				x				x			
10									x				x				x				x				x			
9									x				x				x				x				x			
3									x				x				x				x				x			
0									x				x				x				x				x			
FREQ.	9	31	8	2	20	23	3	4	19	23	7	1	13	25	10	2	15	25	9	1	10	18	16	6	17	21	8	4

bearing on the scalability of the content. The scalogram for the trial intensity is shown in Table 8. Plotting the trial intensity scores against the second trial content scores yields the scattergram in Table 9. Again, the frequencies in boldface in each column indicate the median position for intensity for the respective columns.

As stated previously, 50 cases are far from sufficient to obtain stable column medians when there is a substantial intensity error present, which seems to be the case here. We have strong reason to believe, however, that the absence of a U- or J-shaped curve of medians in

TABLE 9
SCATTERGRAM OF INTENSITY AND CONTENT:
Black Metropolis

INTENSITY	CONTENT (SECOND TRIAL)								TOTAL
	0	1	2	3	4	5	6	7	
21							1		1
20							1	2	3
19						1			1
18					1		1	1	3
17					2	1	1	1	5
16							1	4	5
15				1	2		1	2	6
14	1			1		1	1	1	5
13				1		2	1	2	6
12		1	1				1		3
11				1	1	1	2	1	6
10				1			2		3
9	1								1

3							1		1

0								1	1

TOTAL	2	1	1	5	6	6	14	15	50

Table 9 is not merely due to sampling error, but rather to the fundamental lack of scalability of the content.

An Intensity Curve from a Final Survey. To give the reader a picture of what final results will look like in practice in a complete study, we present some data from a study by the Research Branch.

TABLE 10
AN EXAMPLE OF THE INTENSITY FUNCTION:
JOB SATISFACTION IN THE ARMY

INTENSITY SCORE	CONTENT SCORE											TOTAL
	0	1	2	3	4	5	6	7	8	9	10	
8	23	46	27	33	22	19	24	42	25	23	24	308
7	7	24	31	26	33	31	22	40	21	15	5	255
6	1	7	29	17	30	24	35	42	15	11	..	211
5	6	14	14	29	20	34	27	34	19	10	..	197
4	2	3	15	17	32	33	36	36	10	1	..	185
3	..	1	17	19	22	29	33	25	11	1	..	158
2	1	4	9	19	25	34	31	32	1	4	1	161
1	..	2	2	12	35	39	38	30	8	..	1	167
0	..	3	7	12	29	43	33	26	3	1	..	157

TOTAL	40	104	151	184	248	286	279	307	103	66	31	1,799

Ten questions were asked of a cross section of 1,800 enlisted men with respect to the expression of job satisfaction in the Army. The content was found to be scalable. Intensity was obtained by the two-part technique. The relationship between intensity and content is shown in Table 10. The frequencies in boldface in the columns show the me-

dian intensity for the respective columns. Content score 5 seems to be approximately the zero interval. Men to the right of this score can be said to have positive job satisfaction, and men to the left to have negative job satisfaction.

In conclusion, it should be pointed out that the intensity curve provides not only an objective zero-point, but also a picture of the relative strength with which an attitude or opinion is held. Differing shaped curves, when plotted on the percentile metric, show differing degrees of sharpness in the division of attitudes or opinions. Illustrations of this will be given in future publications.⁷

25. A STATISTICAL APPROACH TO MOTIVATION PROBLEMS*

Some ways by which statistical techniques can be applied to survey data to identify motivations behind consumer purchases are discussed in this article. Regression analysis, the main technique reviewed here, is a well-known statistical tool that has been used infrequently in analysis of survey data. As noted by the author, the technique possesses the advantage of allowing combination of all types of factors—psychological, economic, sociological—in a single prediction equation.

For the last few years, motivation research has been going like a house afire. It hardly seems appropriate to talk about problems of motivation research when motivation research is such a tremendous success in the business community.

Yet, partly because of the very speed with which motivation research has arrived, there are people who have something to do with motivation research for whom it creates problems. One group of people who face a problem are the responsible executives who must decide whether to rely on the findings. One of them may find himself in a situation like this: Here is a report which makes a series of recommendations. Here is an advertising budget of X millions of dollars. Should that money be risked on a campaign built around the report, or should it not? How accurate is that report?

⁷ See L. Guttman and E. A. Suchman, "Intensity and a Zero Point for Attitude Analysis," *American Sociological Review*, Vol. 12 (February, 1947), pp. 57-67.

* Adapted from a talk by J. B. Lansing, Survey Research Center, University of Michigan, "Problems of Motivation Research," given at the American Marketing Association Conference, June, 1954.

Another problem which people connected with motivation research may face is the problem of two reports. Somebody does a piece of motivation research on, say, shoe laces. He comes up with a report which says the important thing about shoe laces is their sensuous feel. But then some unpleasant character turns up an independent report on shoe laces. This report says the important thing about shoe laces is the erotic symbolism of lacing shoes. Sexually thwarted men are very interested in lacing shoes. Nowhere does the second report mention the sensual feel of the laces. Two reports—two contradictory theories. Suppose there were *three* reports!

What is the real problem here? It is, I believe, that too many reports are a mere statement of ideas. Nobody has ever set up a piece of research which really will test whether the ideas are right or wrong.

STAGES OF MOTIVATION RESEARCH

Good research in motivation involves two stages. The informed guessing belongs in stage one. In this stage the purpose should be to collect ideas and insights into people's behavior. The researcher should conduct all kinds of qualitative and exploratory studies; he should cast a wide net to pull in all possible ideas. In stage two he should introduce precise quantitative methods to test the findings of stage one. Stage two is more trouble and more expense than stage one, and it is usually not so much fun. When it is over, there may be no more conclusions than were available after stage one. Very likely there will be fewer conclusions. Hypothesis 1 from the first stage may have stood up like the Rock of Gibraltar, but Hypothesis 2 may have collapsed like a house of cards. Number 3 may turn out to be true under certain conditions only. A brand new conclusion may appear as a surprise. Or it is even possible that not one of the list will have changed. The essential difference between the conclusions of stage one and stage two will be in the strength of the evidence on which the conclusions rest.

The basic question is this: Is it necessary to be right? If not, one can afford to guess. Sometimes a problem isn't big enough to worry about, and the sensible thing to do is to make the best guess one can. But when there is a major problem which has to be solved correctly, the best and most accurate research is none too good.

Let us review briefly some of the techniques which can be used in stage two, with examples from actual research.

RELATING MOTIVES TO BEHAVIOR

The first technique is to measure people's behavior as well as their motives. An investigator need not limit himself to a study of people's

goals, their motives, and so on. He can study them but then go on to relate motives to what people actually do.

In investigations of consumer behavior, it is the buying which is of interest. If particular attitudes are related to whether people buy or not, those attitudes are important. To do that, measures of behavior can be built into the questionnaire along with the measures of motives.

To take a specific example, here is a question about motives, a typical projective question:

Mrs. Jones is a woman who makes long-distance telephone calls to her relatives once or twice a month just to visit with them. Why do you think she does that? What kind of woman is she?

In the same questionnaire the interviewer asked:

Have you or your wife made a long-distance call in the last month? (If yes) About how many long-distance calls would you say you have made in the last three months?

Answers about behavior in making telephone calls can be related to answers to the projective question.

This approach is certainly not new, nor is it peculiar to the Institute for Social Research. Others use this technique. Not every study can relate actual buying to motives, of course, but very frequently this technique can be used in testing the validity of insights and measuring their importance.

REGRESSION ANALYSIS

The second technique is the use of refined statistical methods. Psychologists don't agree about everything, but one point on which agreement exists is the complexity of motivation. A single action by a single individual may be a result of several motives. And individuals differ—the same action by a different person may reflect a slightly different motivational pattern.

Not only do people's motives differ, but their financial positions also differ. People have large incomes or don't have large incomes. They have wives or don't have wives. We have to consider financial and sociological factors.

How can a research worker cope with this complexity? One way is just to form a judgment. We can do better, however.

The statistical tool that comes to one's mind in this sort of predicament is multiple regression. The trouble with multiple regression is it wasn't developed to handle questions which just call for "yes" or "no" answers. It was developed for variables that have a continuous range, like income.

But fortunately it works reasonably well even in situations for which it is not ideally suited. As an example, let me describe briefly some research which uses this kind of an approach, conducted by myself in collaboration with Lawrence Klein. The data are from the 1953 Survey of Consumer Finances, conducted by the Survey Research Center of the University of Michigan in co-operation with the Federal Reserve Board. The object is to explain who of a sample of people bought a car or other major durable item over a period of a year.

For each person in the sample it is known whether he bought or did not buy during the year. Also available is a series of factors that might help to predict whether a person would buy or not. These explanatory variables fall into three groups: financial, demographic, and psychological variables.

The result of the research is an equation. Using the following notation:

B = buying

X_1, X_2, X_3 , etc. = independent variables,

the equation is of the form:

$$B = aX_1 + bX_2 + cX_3 + dX_4 \dots$$

Each of the X 's is tested to see whether it has a coefficient greater than zero, whether the effect is positive or negative, when one takes into account all the other explanatory variables. We have found that common sense is not enough to predict whether a variable will be important.

Let us take the financial variables first—the objective measures of a person's ability to buy. It is not hard to guess that there is a relation between a man's income and whether he buys a car or other major durable goods. But take liquid assets, by which is meant money in the bank and U.S. Government bonds. Contrary to most expectations, it is not true that the more liquid assets a man has at the beginning of a year the more likely he is to buy, if account is taken of the other factors studied.

Would a man heavily in debt be likely to buy durables? It turns out that the more debts a man has at the beginning of the year, the more likely he is to buy. Sometimes one does more psychological research than one intends. Being in debt hardly makes a person financially better able to buy. We interpret this finding to mean that there are psychological differences between debtors and nondebtors waiting to be explored.

Demographic variables have also been introduced in the same rela-

tion. Would you expect buying to be related to marital status? Are married men more likely to buy or not? It turns out that they are. Young people under 25 don't buy as often as older people, even taking into account the effect of marital status. People over 65 don't buy often, either.

After account has been taken of the financial and demographic variables, motivational variables can be added to determine if they improve the relationship. If it is known, for example, that income is related to purchase of a particular product, this information need not be thrown away. That relationship can be taken into account and attempts can be made to improve it by adding motivational or attitudinal variables. In this study it was found, for example, that a feeling of being "better off" financially is related to whether or not a man buys, even after taking into account income and other financial and demographic variables.

What is the effect of expecting an increase in "the prices of the things you buy?" No relation at all was found between expecting an increase in prices and purchasing. But when account was taken of other variables, a relation appeared. People who expect prices to go up are more likely to buy, when you take into account the other variables.

This experience proved again what every statistician knows: One is easily misled in looking at variables one by one.

These functions are very flexible. To plug a variable into a relationship like the one described above, all that is necessary is some technique for classifying people into two or more groups along the dimension of interest. One could, for example, take people who reacted to a Rorschach test one way rather than another, if that was thought to be related to buying some commodity, insert this information into a regression and see whether the relationship improved.

RE-INTERVIEWS

The third technique discussed briefly here is the use of re-interview. It can be combined with the first two—in fact, in the research just described, predictions were based on information obtained in a first interview as measured in a re-interview one year later.

Re-interviews of identical respondents are helpful in several ways. One contribution they make is to solve the problem of circularity. A man, if he has just bought a Ford car, is likely to swear that Ford cars are the best. If he happens to have just bought an RCA television set, he swears that RCA television sets are the best. Or, if he just bought \$2,000 worth of durable goods and you ask him if now is a good time

to buy durables, he will say, "Now is a fine time to buy. I just bought some."

The question is, which came first—his attitudes or his purchases? One way to find out is to measure the attitudes in advance, and then come back for a second interview at the end of a certain period to ask the respondent what he did. Again, this approach can be used at the start in conjunction with any type of measurement of attitudes or motives. All that is needed is to divide people somehow into two groups—those one thinks are more likely to do something, and those one thinks are less likely to do something, and then go back and find out what happens.

AN INTEGRATED APPROACH

Three techniques that can be used to help make motivation research precise have been discussed briefly here. These are not the only techniques nor are they always the best ones. But progress in motivation research requires that, once insights have been obtained, they be put to the test in the most rigorous possible manner.

Why isn't more research done which combines psychological insights with rigorous methodology? There would seem to be at least two reasons. One is a problem of clashing personalities. The kind of research discussed in this article calls for two kinds of skills which are usually found only in two types of people. Type 1 is the man trained in the tradition of clinical psychology or one of the related fields. He is likely to be a perceptive sort of person, sensitive to other people, attuned to what is going on in the other fellow's mind. The chances are that he is not long on quantitative techniques. Type 2 is the statistician, the man trained in mathematics, statistics, and the design of experiments. He is likely to be given to the use of mathematical symbols in his reasoning. Usually he is less sensitive to people and often he is less at home in psychological theory.

These two types of people do not always get along well together. They are sometimes inclined to be critical of each other's work. Hard words are sometimes spoken. The statistician may let slip phrases about the clinician's studies like "interesting—if true," "too subjective," "unscientific," or even "a bunch of guff." The clinical psychologist is likely to talk about the statistician's work as "mountains of dry reports with never an idea in the lot," "tedious elaboration of the obvious," "shallow," "superficial." Too often both are right; there is truth in both types of criticism. These two types of people have to be persuaded to work together more often than is currently the case.

A second reason why there is not more research combining psychological insights with rigorous methodology is lack of sufficient pressure on research people to do careful work. One way of doing this, a technique that will work especially well with academic consultants, is to insist that they publish the research—the problem, techniques, results, and conclusions. Nobody wants to risk his reputation by publishing sloppy research where his colleagues will see it. Of course, competitors will see it, too. But if the competition doesn't know about the findings by the time they are published, the competition isn't very interested.

Let us return to the uneasy executive with whom we started. Though intrigued by motivation research, he hesitated to bet that a particular piece of motivation research was right. I would hesitate, too, if the research had been carried only through the stage of collecting ideas and insights. More research is needed which has been pressed through the second testing stage to the point where the researcher can say to the executive with a clear conscience, "I know that these findings are right and here is the evidence to prove it."

26. WHY TELEVISION COMMERCIALS SUCCEED*

An indirect statistical approach to the explanation of the effectiveness of television commercials is neatly illustrated in this article. In contrast to various other approaches, some psychological, note how the usual formulation-of-hypothesis-then-testing procedures is reversed. Quantitative data are first collected on effect of commercials on product acceptance, the only hypothesis involved being the purely statistical null hypothesis. After these data have been collected and the statistical tests carried out, a search is made for patterns in the observed differences, and only then are explanations of these patterns sought. Such explanations as come to mind are tested against new data collected in much the same manner, just as the adequacy of a regression function can be tested by using it to estimate values not in the period (or space) of observation.

The end-result of such a process can also be a psychological theory of behavior—and the present author comes close to propounding one—but in distinction to results of psychological projective techniques, it is a theory of aggregate behavior and response rather than of individual behavior. In practice, however, this distinction is not often brought out.

* Adapted from a talk by Horace S. Schwerin, President, Schwerin Research Corporation, "Television and Radio Commercials," given at the Marketing Research Conference, University of Michigan, March, 1955.

Television advertising is on the threshold of a new era. We are today seeing the first signs that there soon will be experimentation of a kind and on a scale such as were undreamed of a few years ago. If the trend continues, we can be sure that the commercials of ten years from now will be nothing like the ones we see today.

There are two forces at work to encourage experimentation and change. The first of these is that it has become incredibly costly to remain on television and use an ineffective and imitative sales approach. Advertisers are beginning to realize, in this connection, that relatively little can be done any more in the direction of controlling the cost of reaching viewers, but that a great deal can be done in increasing the effectiveness of the commercials presented to them. Even the shrewdest circulation buy today would give only about a four-to-one advantage over competitors, assuming they made the worst conceivable selection of programs. In contrast, our research reveals that it is possible for commercials for one brand to achieve as much as forty-to-one advantage over another brand in ability to create a preference.

The second force at work encouraging experimentation and change is the development of methods that make it possible for the first time to tell whether new approaches are paying off. The existence of a yardstick like the Competitive Preference technique actually makes it more dangerous for the advertiser to be imitative than to be experimental. The reason for this is that, when his competitor hits upon an imaginative advertising formula, that competitor now has a means of knowing that his new approach will pay off.

The most important single thing that this new measure has brought out is research proof of something that creative people have long hoped was true: Remembrance of copy points is not the whole answer to commercial effectiveness. In the case of conventional commercials, where convincing demonstration is used to put over the right sales idea, there is a close relationship between remembrance and effectiveness.

But that is only part of the story. There is another area besides convincing demonstration, an area which might be called mood or fantasy, that does not necessarily have much to do at all with putting across explicit sales ideas. A commercial of this nature establishes its own world, within which viewers accept actions and breathe in impressions that they would reject if the mood of the commercial were logical rather than emotive. Commercials of this type have proved extraordinarily effective in swaying viewers toward the brand advertised; and

we are receiving more and more commercials of this type to study from advertisers who see which way the wind is blowing.

What has been outlined above might be called "TV's Law of Extremes." In examining the commercials tested that have proved to be most effective, we have found two distinct types. At one end of the curve are commercials where convincing proof of a sales claim is advanced. At the other end of the curve are commercials that create a mood. It is in this second area that boundless opportunities for experiment and progress lie.

Advertisers are basically interested in the cost of effecting an increase in preference; that is, by getting new customers at the lowest possible cost. The way we collectively look at it, there are two distinct areas: First, what does it cost to get into a thousand homes per commercial minute? Second, once a thousand people are captured, how much are the people influenced with commercials, or what is the increase in preference? To provide some perspective on the problem, I would like to cite some costs of getting into a thousand homes.

We use some data borrowed from Mr. Nielsen, based on 98 evening programs on three major networks for the first part of November, 1954, which shows the cost per thousand homes per commercial minute; the best buy was \$1.86 and the poorest, \$7.18, a difference of four to one. We took 95 per cent of the cases in order to get rid of the extremes, where the people are paying excessively high or extremely low rates.

When we consider the second area—what is done once the people are captured—we find an entirely different range. Thus, of commercials studied in 1954, the best commercial had 40 times the effectiveness of the poorest.

The range of efficiency is so startling that, in spite of the fact that circulation and cost are tremendously important, this area seems to be the key to the future in television.

A MEASURING TECHNIQUE

The technique on which the above estimates are based is a laboratory technique.¹ In most of the testing in the greater metropolitan area of New York we attract people from a 40-mile area. We do not go out to them, we invite them into our own theater, located on 46th

¹ Credit is due to the companies who gave us financial support and also loaned us their very fine talent from their research departments, namely, General Mills, RCA Victor Corporation, the Borden Company, Toni, and Miles Laboratories of Elkhart, Indiana.

Street and Sixth Avenue. In Toronto we test at the University theater. In London we plan to test at a small theater on Trafalgar Square. In all areas the techniques are the same.

The people come to the theater. As they walk in, each individual is handed a coupon, divided into two columns, with identical numbers on both sides; one half is retained and the other half dropped into a box. This enables us to identify each individual. We have between 300 and 400 people seated in the theater. Tests are run every night of the week, five nights a week, and three afternoons.

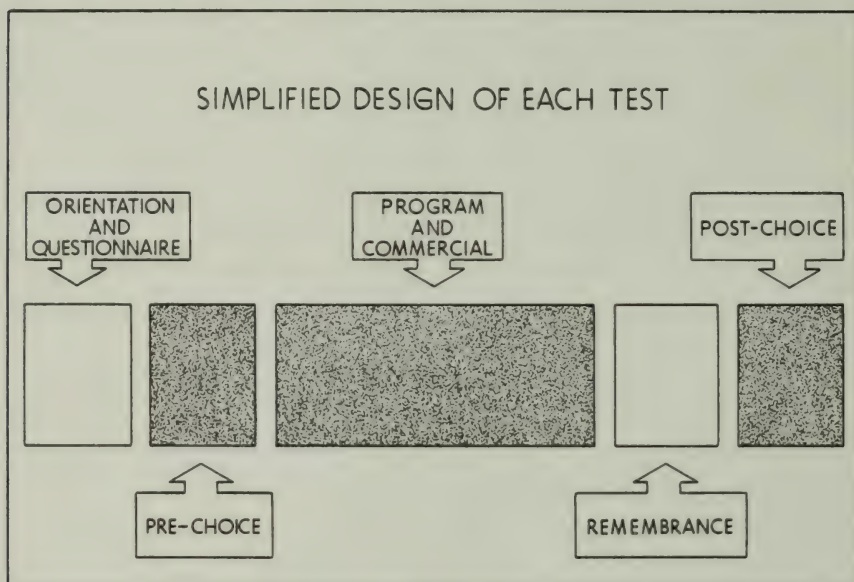


Fig. 1.

This is schematically what happens: First there is the orientation and the questionnaire (see Figure 1). In the orientation a test director shows the people a series of color slides to get them to relax. He emphasizes the point that they are there to help improve programs. Immediately following that, each individual is led through a rather extensive questionnaire, wherein data are gathered concerning personal and socioeconomic classification as well as regarding market information pertinent to whatever the particular subject matter.

Following the orientation and the filling out of the questionnaire, at the point in time labeled "pre-choice" in Figure 1, the people are told that to reward them for having come, a series of door prizes will be given away. Everyone must check, before the winner is drawn, which of the products he wants sent to his home should he prove to be the winner.

For example, the people may have had a chance of winning a Remington electric shaver, a Schick, Sunbeam, or, if they don't want a razor, \$25.00 in cash. They have had an opportunity to win a year's supply of any of these instant coffees: Chase and Sanborn, George Washington, Maxwell House, or Nescafe. They have had an opportunity to win a year's supply of whichever brand of toothpaste they want. Then the drawing follows. The drawing is very simple. It merely consists of a youngster being asked to come up and pull a number out of the basket. The winner is then identified by the test director. At this point the people still don't know what is to follow in the session. At least they have no basis for knowing either that they are going to be seeing commercials or what the commercials are going to contain.

Following the prechoice, there is a half-hour television show with the commercials in the normal position in the show. Immediately following the viewing of the program, the people are asked to play a little game with us. They are handed a sheet of paper and are asked to give the name of the product advertised and everything they remember having seen or heard about the product in the commercial.

Following that the audience is told that we know there were only a few winners previously and that we would like to reward additional people. At this point, there are additional drawings for prizes. They are given fresh forms identical to the ones they had earlier, and they have another opportunity of winning a series of door prizes. Again they make a choice of any of the electric shavers or the \$25.00 in cash. Again there is a drawing. They make a choice of any of the instant coffees, and then a drawing. They make a choice of any of the toothpastes, then a drawing.

If we were studying a particular commercial within a program we would compare what the people wanted before they saw the program with what they wanted after seeing the program and what they remembered immediately following. This group cannot be used for another commercial effort because they are already preconditioned by what they have seen and by their participation in a test. Therefore, to examine this commercial in relationship to another commercial we must use an entirely separate group of people. To compare the results obtained from two or more groups exposed to different commercials, it is critically important that the groups be identical. The research is no better than our ability to make the groups precisely the same with regard to the important factors.

How can this be done? The number of sessions run within the past eight years alone exceeded 2,000 groups of people. According to the

laws of probability, perfect cross sections coming in by accident would be most unlikely.

The means used for obtaining a cross section is postselection. For example, if an equal number of men and women is required and there are 150 men and 250 women, we eliminate 100 women. Of course, that is gross oversimplification, for it is necessary to match on three or four factors and keep them cellularly separate.

Of all the factors that are critically important, by far the most important is the prechoice of the product. Figure 2 shows three separate

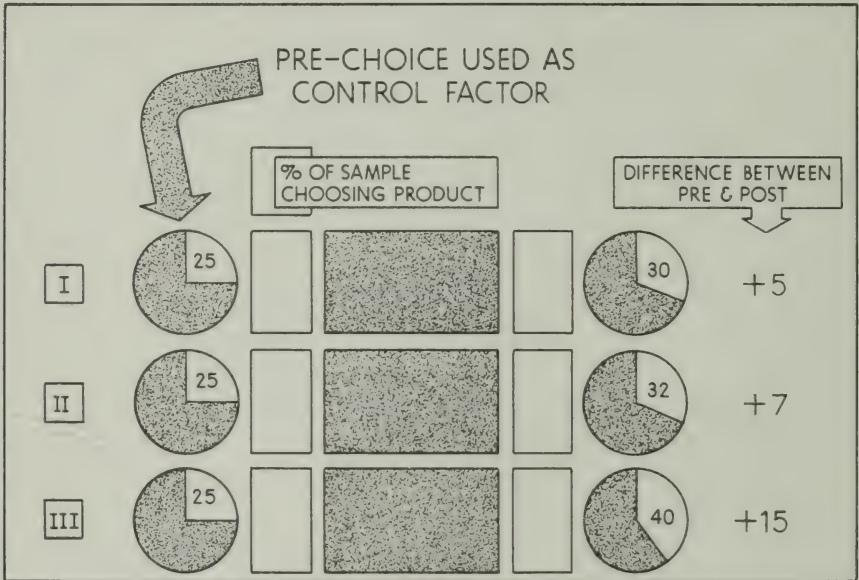


Fig. 2.

groups matched on choice of a hand lotion. To illustrate, 30 per cent of the women wanted this hand lotion in Group 1. That is reduced to 25 per cent, assuming that is the figure selected, by eliminating 5 per cent of the women who wanted it but maintaining all other factors, keeping them cellularly separate. Of Group 2, 20 per cent wanted the hand lotion. That is brought up to 25 per cent by dropping out the people who wanted other brands until the residue is 25 per cent.

This is a very costly method of operation. It not infrequently requires two or three gatherings of people before the residual sample is adequate. In one instance the process of elimination required working six times in order to get the residual sample adequate for a study of Ken-L Ration because of such big differences between dog owners.

Now we have three different groups at the same level. Group 1 started with 25 per cent of them wanting this brand of lotion before

seeing the program. After seeing the commercial in the program, 30 per cent wanted it, an increase of 5 per cent. Group 2, starting with 25 per cent wanting the brand before seeing the commercial, ended with 32 per cent afterwards, an increase of 7 per cent.

Group 3, starting at 25 per cent before the commercial, ended up with 40 per cent, an increase of 15 per cent. We conclude that commercial No. 3 has a stronger influence than commercials 1 or 2, assuming the differences are statistically significant.

RELIABILITY

Let us now examine briefly the reliability of this technique. We are offering the same people at two different points of time the same opportunity of winning a series of products. Do they choose the same when there is no reference to a particular item in between? For example, we offered people cases of Schlitz, Ballantine, Pabst, Miller, Budweiser, Blatz, or \$25.00 in cash. We continued with the regular session, but no reference to beer was made. Then we offered at a second drawing the same list of products or \$25.00 in cash.

In the first instance 21 per cent wanted Schlitz; in the second instance, 22 per cent. Ballantine in the first instance, 17; in the second instance, 17. Pabst in the first instance, 12; in the second instance, 13; Miller in the first instance, 12; in the second instance, 12. Budweiser in the first instance, 10; in the second instance, 11. Blatz in the first instance, 5; in the second, 3. Twenty-five dollars in cash in the first instance, 23; in the second instance, 22.

The changes were not significant. Of course, this had to be done a multiple number of times, as may be illustrated by preference for Schlitz on three different studies on three different groups, wherein there was no commercial in between (see Figure 3). We of course matched to get the same per cent of the people wanting Schlitz, so the groups were identical on brand preference as well as other important characteristics.

In Group 1, 21 per cent wanted Schlitz; no commercial, they ended with 23 per cent. The second group started with 21 per cent wanting Schlitz and ended with 22 per cent wanting Schlitz, no commercial shown. Group 3, 21 per cent, ending with 22 per cent the second time, and no commercial.

We conclude that if we do nothing, nothing happens. This is, of course, an entirely negative type of reliability and evidence is needed on the positive side of reliability. If the same thing is done with groups consistently, something happens.

The same commercial was tested for Toni home permanent waves on three different groups. All groups were matched. At the start, 38 per cent of the women wanted Toni home permanent waves. After the commercial 46 per cent wanted Toni, an increase of 8 per cent. Of the second group, 38 per cent wanted a Toni before the commercial and after the commercial 47 per cent wanted it, an increase of 9 per cent. Of the third group 38 per cent wanted a Toni before the commercial and after seeing the commercial 45 per cent wanted it, an increase of 7 per cent.

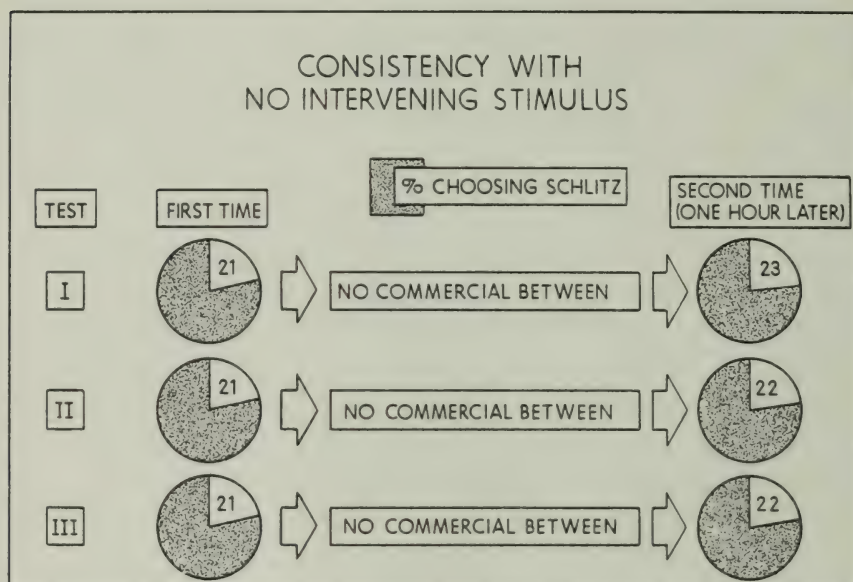


Fig. 3.

Statistically there is no difference between the increments of 7, 8, and 9, so we say that the commercial within the normal limits had the same effect on each of the three groups.

Returning to Schlitz for a moment, we started with the 21 per cent wanting Schlitz. When we did nothing we remained at the same level; 22 per cent wanted Schlitz, or an increase of 1 per cent. On the other hand, if we start another group with 21 per cent wanting Schlitz and they see a commercial for Schlitz and end up with 29 per cent wanting Schlitz, the increase must be attributable to the only change in the experimental structure, which is the inclusion of the commercial.

One might suspect that people would tend to show an increase after seeing the commercials, but here is what actually happens. In the case of a well-known drug product, before the commercial 15 per cent

wanted it, and 14 per cent after; no significant change. In the case of a well-known toothpaste, 11 per cent wanted it before and 12 per cent after; no change. Fifty per cent wanted to win a certain very popular, heavy appliance before the commercial; after the commercial 45 per cent wanted it, a loss of 5 per cent. None of it went to cash; it all went to another make. So it is possible actually for commercials to have a boomerang effect.

A large number of studies have also been conducted in the food field. Of the food product commercials studies, 50 per cent showed no change from prechoice to postchoice that we could measure; that is, the difference from prechoice to postchoice was no greater than if there were no commercial between the prechoice and the postchoice. Some commercials, however, did give increases. Twenty-seven per cent gave an increase of 5 to 10 per cent. Twenty per cent showed an increase from 11 to 20. Three per cent actually gave an increase over 20.

To cite the effects of the commercials that gave increases, the first example is a baking product. Twenty-eight per cent wanted this baking product before the commercial and 32 per cent after, an increase of 4. They were considering a radically new campaign. Before they even tried it on the test market we studied it, with these results: 28 per cent wanted it before and 49 per cent after the commercial, an increase of 21 per cent.

At the time they were running that particular campaign they had a 21 per cent share of the market. Precisely one year later they have a 40 per cent share of the market, without any major change in advertising appropriation.

It is rather interesting to note, in this increase of 21 per cent, that entirely different things are remembered. The campaign that did not do so well had a high remembrance of details, such as the very fine ingredients and the fact that one could bake easily with the ingredients. The campaign that did well had no remembrance of the fine ingredients, no remembrance of anyone being able to bake with ease, speed, and convenience. There was remembrance of good results and the factor of taste.

Let us now consider commercials for television sets. Forty-seven per cent wanted a certain advertised television set in the prechoice and 61 per cent wanted it after the commercial, an increase of 14 per cent. In trying to sell a "clearer picture" idea, this commercial was presented by engineers in white coats, engineers in blue suits, all kinds of professional people, authorities of all types. The campaign that got the highest increase, however, was presented by a baby-sitter. The woman

got up and the first thing she said was, "I know nothing about television." She said, "But I do have to earn my living going from home to home. As soon as the folks leave I quickly turn on the TV set. When I see it is this make I know I am going to get a crystal-clear picture." Interestingly enough, 41 out of 100 people remembered "crystal-clear" and the woman who didn't know anything about television, while 13 out of 100 remembered the engineers' statements.

In beauty products, in spite of the fact that many commercials gave no increment, we have had commercials with an increase as high as 40. One product started with 20 and went to 60. In the appliance field quite a few have done nothing, yet some commercials have given as high as 25. In the drug field quite a few also gave no response. The highest has been an 11 per cent gain. In soaps and cleansers, many commercials have given nothing. The highest we have had is 10.

WHY EFFECTIVENESS VARIES

In showing commercials and studying the increments, the highest effectiveness was obtained the more clearly the product was demonstrated, that is, when the public had an opportunity of using their own eyes to see the payoff idea, such as with Hazel Bishop lipstick—it won't smear off (and they show it); it won't stick on your cigarette (and they show it not sticking on the cigarette). Where the demonstration is one in which the people are seeing the payoff of the product, invariably, in terms of effectiveness, we are way up on the scale.

With commercials farther down the scale people are no longer being allowed to use their own eyes; somebody is doing it for them. Well-known authorities are saying, "This is the best product." In no instance does that do as well as where people can use their own eyes. Still farther down the scale are commercials in which somebody who is not even an authority stands behind a desk and hammers on the desk, assuring people that his product is best.

We found in one case that the lower the belief, the lower the effectiveness; the higher the belief, the higher the effectiveness. The lower the remembrance, the lower the effectiveness; and the higher the remembrance, the higher the effectiveness. So there are three factors correlated: The higher the remembrance and the higher the belief, the higher the effectiveness, and the difference between the pre- and the post-increment.

Fortunately, the same basic product was presented in a variety of ways. For example, let us take soap. A shampoo was presented in various ways. An actual demonstration was first, showing the payoff

idea of the product or the motivating idea of the product by live demonstration.

In a simulated demonstration, a cartoon or some other device is substituted for the live demonstration of the payoff of the product. In an illustration of results, the picture is shown of a fat woman before using the product, and then a picture of a thin woman is shown. That is indicative of what we mean by results. In direct testimony somebody stands up and says, "Trust me. This is the best product in the world."

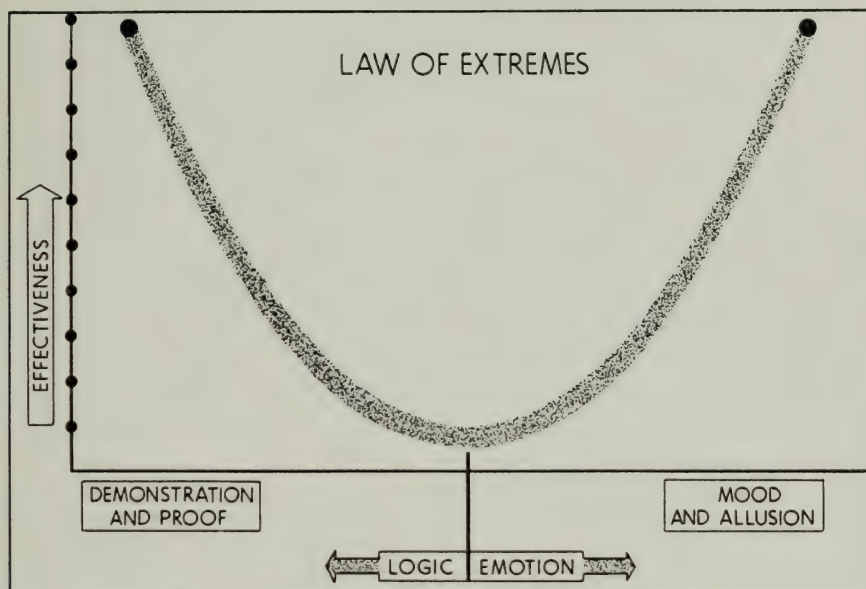


Fig. 4.

In indirect testimony, we could not figure out why the commercial was produced.

I would like to point out that where the actual demonstration took place the cluster of the highest range appears. Where commercials tended to go into other approaches they did not do as well.

The demonstration in itself is subject to variables. For instance, one model is shown getting the best possible suds when she uses a certain shampoo. Then there is a model who is making a comparative demonstration. This model uses a common, ordinary shampoo and she doesn't get suds. The difference is rather interesting. Showing the one model alone using the product, we got 20 ideas back to 100 people. When we showed a rival product with two models, we got 40.

Most of the commercials we are talking about (see Figure 4) were essentially radio commercials to which video was added. The visual

was not used as a prime means. We then move into another area that nobody is expected to believe. It is obvious that it is fantasy. The farther we get into the mood and illusion area, the greater the increment. The single highest scoring commercial in terms of effectiveness tested so far had only 30 ideas remembered out of each 100, which is about one-sixth of normal for that product. It had no belief because nobody was expected to believe the particular fantasy shown. The only thing that happened was that women sure wanted the product!

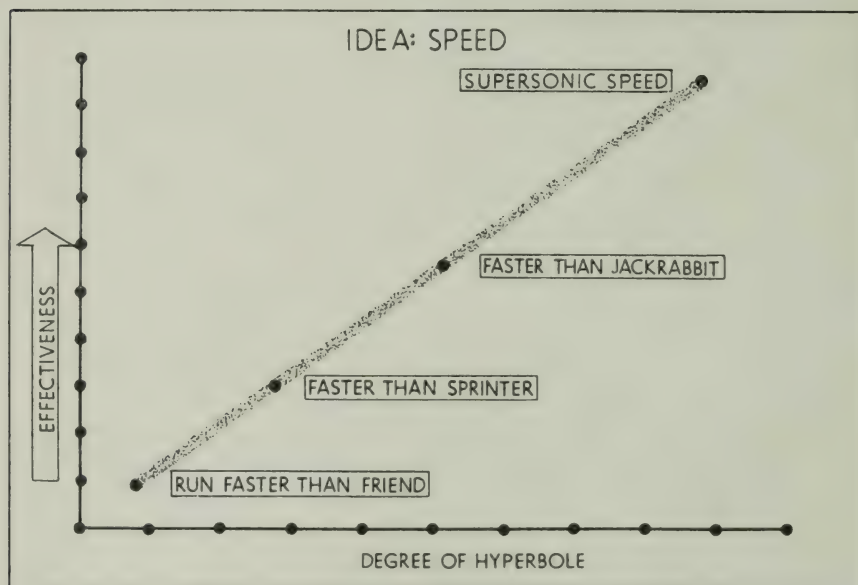


Fig. 5.

Indication of effectiveness of commercials in this area of emotion is a commercial for sneakers (see Figure 5). The vertical scale is the effectiveness, and the horizontal the degree of hyperbole. The point of one commercial was that the boy can run faster than his friend because he has sneakers on. We get low interest.

As a result of wearing the sneakers he is faster than a sprinter. We get higher interest. When as a result of wearing the sneakers the boy can travel at supersonic speed we get even higher effectiveness.

This may seem silly, yet there are many other examples. Up to very recently most advertising people have been concentrating on getting across more ideas, having to prove them, and having to demonstrate them. For some reason they seem to forget that there is more than one way of influencing people. Some advertising is capable of arousing the desired reaction. The very best commercials—and unfortunately they

are very far and few between—are commercials which are in the area of “mood,” or whatever one may want to call it. Most commercials in past years have jumped from logic to emotion. The mood of the program plays a large part, particularly with commercials which are selling via mood, emotion, illusion, allusion, whatever is the correct word.

PART IV

Other Techniques

INTRODUCTION

As some of the preceding selections have already noted, not all problems of motivation require surveys as the method of solution. In fact, important clues, if not complete solutions, to many such problems can be obtained at times through analysis and examination of secondary data. The techniques used in such analyses vary considerably in both range and complexity. The selections reproduced in this part of the book are designed, like those in Part III, to provide some idea of this variety. Articles such as those by Harvey and by Zipf also serve to indicate that even primary data collected by means other than surveys of human populations can be of use in motivation studies.

At this point, however, some readers may well ask: In what sense can studies such as those reproduced in this part answer questions of human *motivation*? There are two answers to this query. First, analysis of the end-result of human behavior can yield, at times, more reliable information regarding the psychological bases for this behavior than can actual interviews—and invariably at considerably lower cost. Thus, comparison of the content of best-selling books with that of books that did not prove popular uncovers the type of appeal that proved successful, which in turn can be translated into the psychological factors that motivated people to buy one type of book rather than another type. Final proof may well require additional work designed to test the predictive accuracy of the hypotheses developed from such a study, although even a limited study of this nature can be valuable enough to book publishers.

Second, as was noted earlier, there are many dimensions of motivation, not all of which are psychological, and in numerous instances psychological motivations are not necessarily the most relevant for marketing policy. This is well illustrated by Steele's analysis of the effect of weather on the sales of a department store. Not surprisingly, weather is found to influence such sales substantially. Hence, the reason why a department store's sales may be somewhat lower one time than another may be due entirely to natural influences, and it is only after allowing for such influences that one can proceed to determine if other forces (psychological or otherwise) might also be instrumental.

This example points out once again that in many problems the question of "why" and the question of "how much" are virtually inseparable.

In the course of answering one question, the other is answered also—and where possible this joint-solution approach clearly is most preferable for marketing policy.

Furthermore, this and other examples in this part provide many indications of how psychological and other approaches can, and should, be used jointly. Usually, the former is applied first to obtain clues to behavior and working hypotheses. Nonpsychological techniques ranging from surveys to time series analysis then serve to test the validity of these working hypotheses. In some cases, a series of such steps may be needed in iterative fashion until a final solution is reached. In any event, it is clear that the researcher who relies on only one type of approach to solve problems of market behavior is doing a disservice both to himself and to those relying on him.

27. WHY CONSUMER PURCHASES CHANGE OVER TIME*

One of the most basic problems confronting business researchers is the prediction of consumer purchases. But a prerequisite to accurate prediction entails, as Ruth Mack aptly points out, prior understanding of why purchases are made at some times and not at others. In this article, a broad theoretical framework is advanced as one means of dealing with this problem and is applied to an explanation of the variation in shoe sales over time.

Though the method is based on analysis of aggregates, it is interesting to note how working hypotheses are derived from considerations of individual behavior.

Consumer buying is a function of the nature of consumers, on the one hand, and the nature of buying—that is, the decision to buy—on the other. Let us consider each briefly in turn.

CONSUMPTION IN OUR CULTURE

Buying or rather the result of buying—consumption or possessions—constitutes a very large and very important part of the interests, values,

* Adapted from a talk by Ruth P. Mack, National Bureau of Economic Research, "Analyzing Changes in Consumers' Buying," given at the meeting of the American Marketing Association, December, 1953. It is based on material published in *Factors Influencing Consumption: An Experimental Analysis of Shoe Buying*, Technical Paper 10, National Bureau of Economic Research. A shorter version of this paper constituted a chapter in *Consumption and Business Fluctuations: A Case Study of the Shoe, Leather, Hide Sequence* (New York: National Bureau of Economic Research, 1956).

and actions of most people. We tend to forget this very significant fact because, like most significant facts, it is all around us and therefore nowhere. But all one has to do to realize its importance is to glance through descriptions of primitive cultures or even of oriental cultures. These are likely to convey a startling awareness of our own cultural patterns. Vividly it is borne upon one that, whereas some peoples' life may center on ancestor worship, propitiation of fertility (vegetable or animal), warlike conduct, ours centers around buying—or rather the results of buying—possession and consumption. Much of our basic sense of importance, our ability to measure up, our attitudes toward others, roots in the things that we own and use, and this fact is of enormous importance in analyzing consumer buying for several reasons.

First, it must mean that there is a tight two-way bond between consumer disposable income and buying. The desire to use and to possess goods drives people to earn money (or drives wives to drive men to earn money). The intensity of this drive is a cultural matter.

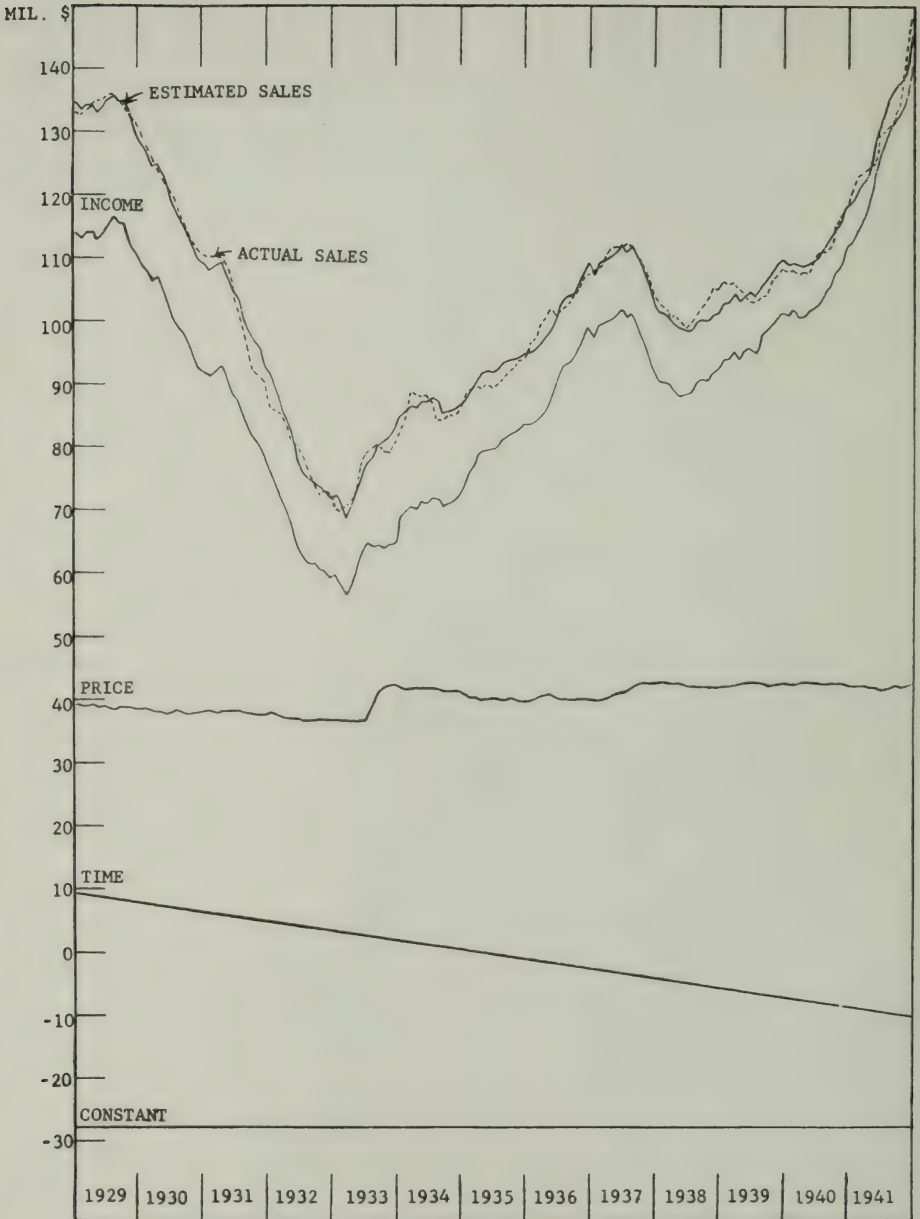
Second, since urgent wants are virtually limitless, spending tends to hug the ceiling on the ability to spend. The ability to spend is limited by current income, total savings, and the ability to borrow. Since most American families do not have large quantities of uncommitted savings, and their ability to borrow has fairly narrow limits, the size of current income will inevitably be the major limiting factor in current buying. In a culture in which wants are intense, central, and virtually boundless, the limiting factor—income—becomes the controlling factor in major change, as is evident in Chart 1. Unlike the familiar association of total consumption and total income, two of the time series in the chart—the one for shoe sales and for income—have no single common element in their statistical derivation, and yet we see a sensitive parallelism between the buying of this one commodity, shoes, and monthly disposable consumer income.

THE NATURE OF BUYING

The second theorem for analyzing buying to be derived from the basic importance of consumption in Western culture is that, though income is the major determinant, virtually anything else at all can influence what people buy. For virtually anything that happens to people will have some reflection in those things which are most important to them—an unconscious as well as conscious reflection. In many cases, the things that occur, such as changes in expectations, in the size of families, in the owned stock of usable goods, will have a direct and logical bearing on buying decisions. This means that the analyst who

CHART 1

CONTRIBUTION OF EACH OF THREE VARIABLES TO THE ESTIMATION
OF SHOE SALES, 1929-41



tries to predict consumer purchases must be prepared to see almost anything move his time series up or down on the paper.

The power of some factor to produce an important impact on buying will be a function of the extent to which buying is innately sensitive

to it on the one hand, and, on the other hand, of the amount that the factor has itself undergone change. But the power to observe and measure its influence will be a function of the extent to which the time pattern of the factor in question differs palpably from that of other important variables. We shall come back to these matters later.

A third theorem that may be derived from the basic notion of the importance of consumption in our living pattern, and the associated drive to make money, is the omnipresence of time trends in the absolute and relative buying of individual commodities. For the drive to make money (in order to consume or whatever) means that new products are constantly being invented and developed, and the drive to consume means that place must be made in the budget for buying these new articles, especially when (as in the case of television, radio, and, long ago, phonographs) special kudos is attached to their ownership. This means that the purchase of other articles must give way as the new stars rise. It is not at all unlikely that, other things the same, this downward trend relative to income is found even in the purchase of that complex bundle of utilities provided by the saved dollar. The point is that economists need not apologize for using (as a rough approximation to something real) a time trend in regression analyses of consumer buying. Note that shoes were, as indicated in Chart 1, one of the commodities that experienced this downward trend in the proportion of the income dollars devoted to their purchase. Many things contributed to this trend, both positively and negatively, besides the general one already mentioned.

To continue with this chart, we have thus far learned something about why the line for shoe buying follows income so closely (and why, also, it would not have to—a paradox to which we must return). We have learned that, unless price developments are unfavorable, the buying of this one ancient commodity may well decline relative to income over the years. But as yet we have not learned why for every dollar of change in income, shoe buying seems to have changed by a certain amount— $1\frac{2}{3}$ cents—whereas out of each dollar of income, a bit under 2 cents of average income is spent on shoes. Obviously both these figures would be very different for other commodities. Why? To answer this question, it is necessary to inquire into the character of a decision to make a purchase, a complex subject.

DIMENSIONS OF BUYING BEHAVIOR

The decision to buy is a thing of many dimensions, some of which are peculiar to each sort of article bought. But at least four of these

dimensions are basic to most of buying. The first involves which needs are filled. Since wants are almost infinite and means all too finite, it is necessary to select among needs by rationing the scarce means so as to maximize total satisfaction. This rationing has a basic pattern of far-reaching importance. It is hierarchical—some needs are more important than others and at least partial satisfaction of more important needs is prerequisite to interest in less important ones.

The second dimension of a decision involves the actual article selected for satisfying the want. In this process, price plays its much publicized role. When the want is specific and only one article will clearly satisfy it—the favored brand of cigarettes is a clear example—this dimension merges into the first. When the need is specific and yet several articles might satisfy it, a choice among candidates must be made. The purchase of various sorts of foods to satisfy hunger is an example of this sort of selection and one in which the knowledge necessary to choose the appropriate article is typically at hand. Often it is not, as in the case of women's cosmetics, though the need is quite clear—the fountain of Ponce de Leon.

But it is possible, and indeed usual in a high level economy, for the need itself to be both unclear and complex. Most food is purchased not only to satisfy hunger but to be enjoyed for its taste or for its power to yield a sense of well-being. Yet the prime example of this complex bundle of wants that a given article may fill is the status purchase. On which article the mantle of status falls is often highly conventionalized, as in the case of the television set. Sometimes it is an individual matter.¹

Needless to say, in the creation of wants, their place in the hierarchy, and the definition of what articles will satisfy the wants, our society is changeable. In part, this mobility is a concomitant of growth and mobility elsewhere. In part it is directly induced or accelerated by movies, radios, television, public relations councils, and advertising experts of all sorts.

If we put these two dimensions together—the hierarchy of wants and the articles that satisfy them—and connect them by the web of price, we have in effect a hierarchy in the relative use-value of various articles. But this order in use-value cannot be translated into an order in purchasing until the relationship between use and purchase is defined. For obviously, the relationship varies widely among articles.

¹ I know of two families in Vermont, neither one of whom could have had \$800 in cash in their hands that year, one of whom spent \$70 on Revere pots and the other \$60 on a red toy automobile. It is not at all clear that these were ill-advised purchases, for they had a complex symbolic value.

This brings us to the third dimension of a purchasing decision—when to buy. It is probably this dimension that is most elaborately involved in business fluctuations. The fact that many large, more-or-less durable articles continue to give service long after they are purchased makes it possible to postpone purchase when income is low and invest in these articles when income and ability to borrow is high. The result for the economy as a whole is, of course, a far more violent swing in the buying of durable than of perishable goods. One characteristic of the postponable decision is that it often takes considerable time to mature. One reason why we know so little about purchasing of durable goods is because we have not given sufficient attention to the way in which a decision to buy an automobile or a television set typically starts, gathers force and finally—and only finally—results in a purchase.

The fourth dimension involves price. The relative price of articles offered for sale plays an intricate and oft-described part in this whole buying procedure. No aspect of economics is more thoroughly explored, at least theoretically. But we must also be concerned with the several qualities, and consequently prices, at which a given sort of article may be purchased. For it is characteristic of the articles offered for sale that they are typically available at different price lines which range from the strip model to the superior special de luxe model with all the extra gadgets. The average price paid for most articles usually rises with family income, just as the number of at least most of these articles purchased typically rises. The relative importance, however, of the change in the number as compared with the change in the price in contributing to the change in total expenditure differs very markedly among articles.

INCOME AND SHOE PURCHASES

In the case of shoes, for example, this general pattern is evident in the 1935–36 Consumer Expenditure Survey. By and large, both the price paid for shoes and the number of pairs bought rise with income. But there is a difference between men's and women's shoes. Men do not buy many more pairs of shoes as their income bracket moves up, but they pay considerably more for each pair. Among women, the desire to buy more shoes keeps pace with the desire for better quality. All this seems very reasonable, but it is strange that it has been so little studied. It means that cyclical increases in income are likely to be associated with "trading up" on the part of the consumer and cyclical declines with willingness to accede to the notion "it's smart to be thrifty." This fourth dimension of the purchasing decision is very important if we

want to analyse the number of various articles purchased (as contrasted with their aggregate dollar value). It introduces interesting technical questions in adjustment for price change and comparing retail sales (typically reported in dollars) with manufacturing output (typically reported in physical units).

But it is the other three dimensions that are involved in the question with which we started this discussion of the buying decision—why shoe sales rose or fell by about $1\frac{2}{3}$ cents for every dollar's rise or fall in income, and why they constituted a bit less than 2 per cent of average income. The difference between the average and marginal propensities to consume shoes is a function of the fact that the observed relationship between shoe buying and income is such that were income zero, shoe buying would still be sizable. (The true constant in the charted equation includes the minimum price and trend values.) In other words, some minimum amount of footwear falls at the base of the hierarchy of articles for a substantial number of people.

The marginal propensity to buy shoes, on the other hand, is about 1.65 cents per dollar of average income, which corresponds to an income-elasticity figure of about .8; a typical marginal propensity for consumption as a whole. This suggests that shoe buying is more or less central in the array of uses to which new income dollars are put.

Other commodities differ drastically with respect to their place in the hierarchy of wants. Luxury price lines of articles tend to have high marginal propensities, along with durable goods the purchase of which can be put off to a time when income is high and the willingness and ability to borrow strong. Staples, such as some foods, housing, and drugs, have very low propensities. This tendency for more of newly gained or lost income dollars to be spent on some commodities than on others means that, when business and income improve, industries producing income-sensitive commodities will gain relative to those producing the less income-sensitive ones, just as they lose more when income falls.²

² The possible impact of this differential change can be judged from a set of calculations made a few years ago at the U.S. Department of Commerce by Fred Winston and Mabel Smith. By means of regression analysis of dollar consumer expenditure, well over 100 commodity and service groups were classified into those having average, greater or less than average income-sensitivity. They ranged from an elasticity of over 3 for yachts and pleasure crafts to .12 for gas and electricity. Unfortunately, differential price change was not eliminated from the picture. Comparing the proportion of total sales consisting of goods above average sensitivity, in prosperity and depression, we find that for two pre-World War II peak years, 1929 and 1937, the figure was 27 per cent and for two depression years, 1932 and 1933, 21 per cent. The corresponding figures for goods below average sensitivity were 36 and 45. This reflects very different histories for stable and cycle-sensitive consumer goods industries and involves a fairly considerable shift in the composition of the national product.

In many ways, then, the several dimensions of a buying decision affect the relationship of purchases to the major determinant of buying—income. But, since use and ownership of goods is so central a cultural value, a person's buying is a fundamental expression of his personality and can be affected by anything that affects personality. This is virtually anything that has impinged on the individual—past or present or expected future. Are there elements in the current scene that might have a substantial influence on buying? If such factors have themselves undergone substantial change which can be isolated in a time series, such influences might be both significant in influencing the pattern of that buying and subject to rough measurement.

Again using shoe sales as an example, let us return to Chart 1. In addition to income and time, price seems to have a systematic influence on buying. This is a function of the need to choose among alternative uses for money and the fact that, if prices of a given good increase relative to others, goods foregone in order to purchase it increase; the greater sacrifice tends to decrease buying of the good whose relative price has risen, other things the same. Though there are reasons to mistrust the calculations, they suggest that, if prices went up one per cent, the physical volume of shoe buying would be reduced about a half of one per cent. Other commodities would be expected to have a greater and others a considerably smaller price-elasticity of buying.

FACTORS OTHER THAN INCOME

On the basis of logic, introspection, and some preliminary evidence, other factors also seem capable of imparting a systematic bias to the buying of different commodities at different times. Take the question of whether income is rising or falling—the rate and direction of change in income. Some expenditures are notoriously difficult to adjust to recent alterations in income. Rents are an example; their adjustment is restricted because moving is itself expensive and full of social significance. Small purchases made by a considerable number of people in the family are another example. These things might show some lethargy in the speed with which they fall in line as income changes, so that, other things the same, expenditures on them would be expected to be higher had income recently fallen to a given level than had it risen to that level. For goods the purchase of which is motivated partly by future prospects, the reverse might be true; durable goods, for which the timing of the purchase has a strong element of election, are cases in point.

Expectations are another factor that may exert an important influence

on buying. Expectations concern income, price change, or the availability of goods. Many people are likely to undergo similar changes in expectation at a given time. A dramatic example of how buying can be influenced by such factors is provided by the spurt of durable goods purchases associated with the Korean war in 1950–51. Another example of the impact of expectations, this time concerning prices, seems to have occurred at the time of the Blue Eagle and NRA in 1933; it is evidenced in sales of certain departments of department stores.

The impact of stocks of commodities that consumers already hold also may influence current buying. This is an exceedingly tricky matter.³ For it follows from the relation of buying to other cultural values that, though people's current buying may be lower when their possessions are high, their current buying may also be higher when the possessions of their friends and neighbors are high, other things being the same—man is a social animal. Thus, the inverse association to one's own stocks and the direct one to the stocks of others will often cancel out; or one or the other may be quantitatively more important.

Because wants form a hierarchy, the buying of any one commodity can be positively or negatively associated with that of one or more other commodities. Purchase of furniture is positively associated with housing purchases and housing purchases with purchases made by wives; money saving, particularly in the postwar period, is negatively associated with durable goods buying; and so on.

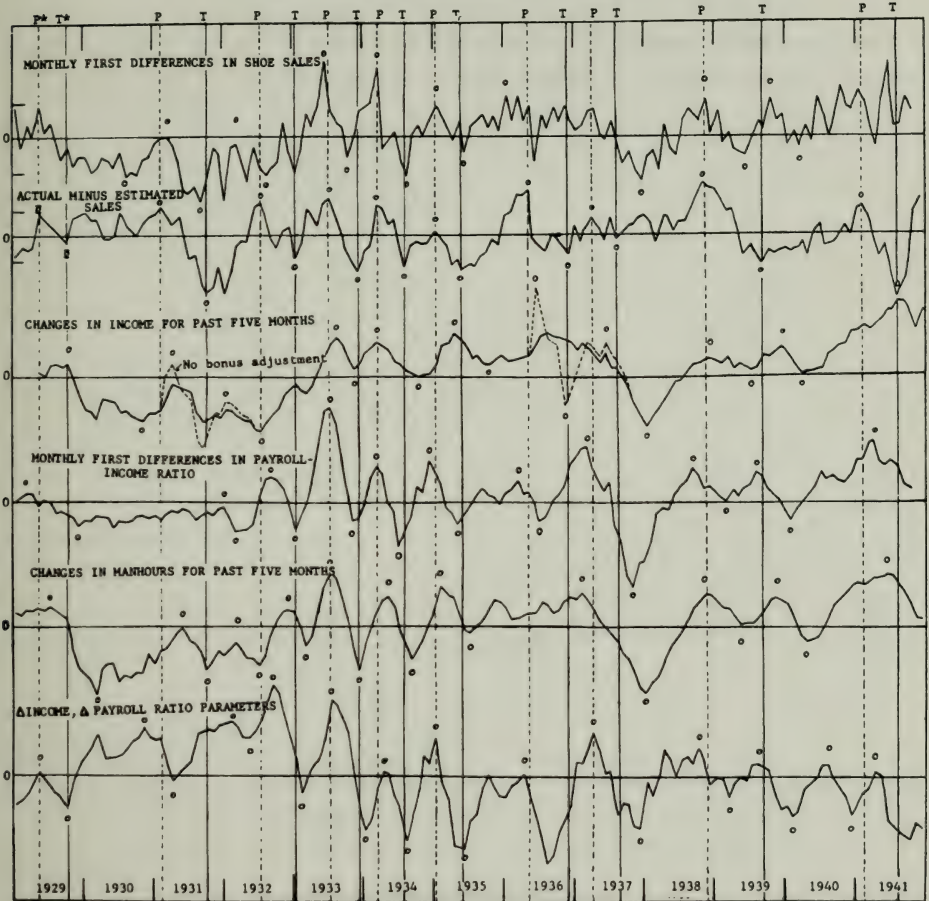
The extent to which any or all of these possible influences help to explain the course of shoe buying during the 15 years before World War II is shown in Chart 2. This chart is a cryptic statement of matters developed at length in the monograph previously mentioned. It starts with the element in shoe buying unexplained by the three variables—income, time, and relative price. This second line on the chart is simply actual sales minus those computed in the regression formula. By comparing it with the top line, we see that this calculation shared a usual attribute of regression calculations, that of failing to explain the rate of change in the dependent variable. The unexplained residuals are compared with time series representing some of the factors just mentioned. Though the sign of the association seemed to accord with expectations in several cases, the influence was too small to be trustworthy. The difficulty may well be a technical one: during the period studied, income varied by huge amounts. Income is not only an extremely important influence on buying in its own right, but it has a way of paralleling

³ It is discussed at some length in the reference listed in the opening footnote.

time series representing some of these other factors that we have mentioned. As a result, these lesser influences may not be visible. At a time, such as the past twelve years, when income has changed relatively little but other factors have changed considerably, the influence of income may be the difficult one to isolate.

CHART 2

ERROR OF ESTIMATING SHOE SALES BY 3-VARIABLE EQUATION COMPARED WITH INFLUENCE OF ADDITIONAL PARAMETERS



*Peaks and troughs in actual minus estimated sales.

Even over the prewar period, the influence on shoe buying of factors such as expectations may be visible in the chart. If so, they are seen at times when they were especially important, rather than in a systematic fashion throughout the stretch of years. The impact of anticipation of price change is suggested by the unexplained spurt in shoe buying in June and July, 1933, just before the big increases in prices actually appeared. In early 1936, buying increased in anticipation of the distribu-

tion of the soldiers' bonus bonds in May; in 1938, the unexplained height of shoe buying might have been attributable to real style news in the shoe field. These points cannot be pushed too far, but they suggest that it is wise to consider not only quantitatively measurable influences on consumption, but qualitative ones as well. In this fashion, exceptional as well as regular influences can be taken into account.

If it is true, then, that in this time and country consumption is so nearly co-extensive with personality, and buying decisions are really such multifaceted things, then to view buying properly we must view it broadly—and, especially, flexibly. For whenever some quasi-environmental factor bears on many people in the same way at the same time, their buying is likely to respond in an appropriate fashion. To these responses the ear of the analyst must be attuned as well as to the familiar and sometimes deafening roar of the rising and falling tide of income.

28. WHY ADVERTISING READERSHIP VARIES*

Why some advertisements receive more attention than others has long been one of the key research problems in the advertising field. Research in this area has generally followed one or the other of two broad approaches—intensive interviewing of readers followed by an attempt to relate readership to personal characteristics of the respondents, or correlation of aggregate readership data with distinguishing characteristics of the advertisements. The present study is a good example of the latter approach, showing among other things how the psychological-statistical technique of factor analysis can be employed in this type of work. As with other studies of this type, the results do not throw much light on people's "inner" reading motivations. For business magazines, however, the approach used here is likely to yield even more meaningful results than might otherwise be possible, for the reasons noted by the author.

One of the most widely used indices of the attention value of published advertisements is the extent to which people read and remember, as determined by readership recognition surveys of the Gallup, Starch, and Advertising Research Foundation type.¹ In these surveys, a rep-

* Adapted from an article by Dik Warren Twedt, Leo Burnett Co., Inc., "A Multiple Factor Analysis of Advertising Readership," *Journal of Applied Psychology*, Vol. 36 (June, 1952), pp. 207–15.

¹ Basic data for the present analysis are taken from the Advertising Research Foundation's Continuing Studies of Readership. The ARF, a nonprofit organization sponsored jointly by the American Association of Advertising Agencies and the Association of

representative sample of a publication's circulation (usually from 200 to 400 subjects) is interviewed shortly after publication of the survey issue. Working with a whole copy of the issue (or an abbreviated issue if the original is so large as to cause fatigue during the interview), the interviewer goes through the issue page-by-page, recording the elements which the respondent says he has read. The resulting readership scores are simply percentages of readers who report having read a particular article or advertisement.

The present analysis is based primarily upon the Advertising Research Foundation's Continuing Studies of business magazines² for these reasons:

1. They are the most recent studies published by ARF, and have advantages of certain technical refinements, such as Lucas' confusion-control,³ the elimination of respondents who identify more than a critical number of advertisements or articles which have never been published.
2. Business papers which are members of the Audit Bureau of Circulations (as are all of the publications represented in the business magazine studies) are required to classify their subscribers by occupation and geographical area (see paragraph 10 of the biannual publisher's statement, available from the Audit Bureau of Circulations or from the publisher).
3. It is reasonable to assume that the population of readers of a business magazine such as *American Builder* is more homogeneous with respect to interest in business advertising, than are readers of general media with respect to consumer advertising.
4. In measuring readership of consumer advertising in general magazines, it is difficult to partial out the cumulative impression of advertising in other media, such as radio, television, and billboards. This problem is also present in business paper advertising, but to a considerably lesser degree.

PURPOSE OF THE ANALYSIS

This analysis has a threefold purpose:

1. To define and measure certain variables in business magazine advertising, and determine the interrelations among these variables, and their relation to readership as measured by the ARF recognition surveys.

National Advertisers, has as its purpose the promotion of greater effectiveness in advertising through impartial research. The foundation makes newspaper, farm paper, transportation advertising, business magazine, and executive management publication readership studies. Since its inception in 1936, the Foundation has published 180 surveys of nine media in 146 markets throughout the United States and Canada. See Advertising Research Foundation, *The Advertising Research Foundation: What It Is—What It Does* (New York: Advertising Research Foundation, 1951).

² Advertising Research Foundation, *Continuing Study of Business Papers*: No. 1, *Automotive Industries*, issue of October 15, 1948; No. 2, *American Builder*, issue of February, 1950; No. 3, *American Machinist*, issue of March 6, 1950; No. 4, *Chemical Engineering*, issue of March, 1950 (New York: American Research Foundation).

³ D. B. Lucas, "A Rigid Technique for Measuring the Impression Values of Specific Magazine Advertisements," *Journal of Applied Psychology*, Vol. 24 (December, 1940), pp. 778-90.

2. To determine the factorial structure of the relationships among these variables, so as to make possible a simpler psychological explanation of the obtained variance in readership scores of advertisements.
3. To develop a multiple-regression equation which will predict advertising readership in business papers.

This analysis is thus one of audience (or what people do to the advertisements) rather than one of effect (what the advertisements do to people). The same general experimental and statistical approach is also applicable to studies of advertising effect, the only stipulation being that an adequate effect criterion must first be available.

The experimental design employed in this study is intended to uncover general principles of advertising which will increase the probability that prospects will be exposed to a given sales message. Because of the complex nature of the problem—the many variables which may influence readership both directly and through interaction with other confounding variables—and particularly because of the expense and difficulty of controlled, single-variable experimentation in a practical advertising situation, it is not easy to evaluate the relative importance of variables contributing to variance in readership scores. Comparison of high-scoring advertisements with low-scoring advertisements is helpful, but this does not represent the most powerful statistical technique available. And we do need statistical controls; even where large numbers of observations are available, categorizing the data by such pertinent variables as size and color may reduce the number of cases so greatly that conclusions based upon them are not likely to be stable.

Fortunately there is an exploratory method (multiple factor analysis) which is well suited to the Continuing Studies of readership data. The basic assumption of the factorial method is that there is an underlying order which, when found, will permit us to give a simpler explanation of phenomena which may seem to be the result of a very large number of variables. The method begins with a table of intercorrelations, or "correlation matrix" (see Table 2). From this matrix we attempt to get simplified explanations or factors for the observed correlations (see Table 3).

PROCEDURE

A preliminary analysis was made of the ARF Continuing Study of Business Papers No. 2,⁴ on the *American Builder*, a monthly trade magazine edited primarily for building contractors and dealers. At the

⁴ Advertising Research Foundation, *Continuing Study of Business Papers*, No. 2, *op. cit.*

time of the February, 1950 issue, its circulation was approximately 80,000.⁵ The *American Builder* was chosen for this analysis principally because of willing co-operation from the magazine's publisher and research manager.

This magazine averages more than 300 pages to an issue. The survey issue of February, 1950 contained 320 pages, of which only 188 were included in the restapled interviewing copies. The abbreviated survey issue contained 137 advertisements of varying sizes, ranging from $\frac{1}{8}$ page to 4 pages. In advertisements $\frac{1}{4}$ page or larger ($N = 122$), the following readership percentages are available:

1. "Any This Ad," per cent who remembered reading or seeing any part of the advertisement.
2. "Headline," per cent who remembered reading the principal headline of the advertisement.
3. "Any Copy," per cent who remembered reading any of the advertising copy, exclusive of the headlines.
4. "Pictures," per cent who remembered seeing the picture indicated.

For advertisements smaller than $\frac{1}{4}$ page, only one readership percentage, "Any This Ad," is given. For all advertisements $\frac{1}{4}$ page or larger, "Any This Ad" readership percentages correlated .98 with "Pictures," .91 with "Any Copy," and .90 with readership of "Headlines." The more inclusive category, "Any This Ad," was chosen as the criterion measure of readership.

Against this criterion, product-moment r 's were computed for 34 advertising variables (see Table 1). Mechanical variables are listed as items 1 through 15 in Table 1, and content variables are listed as items 16 through 34.⁶

The correlations of .00 and .01 between readership and Flesch readability indices⁷ are not statistical artifacts due to restriction in range of Flesch scores, but they may be a function of high specialization of interest by technical audiences.

Of the 34 variables which were correlated with the readership cri-

⁵ *Standard Rate and Data Service*, Business Publication Section, Vol. 32 (1950), p. 116.

⁶ Detailed definitions have been deposited with the American Documentation Institute. Order Document 3417 from American Documentation Institute, 1719 N St., N.W., Washington 6, D.C., remitting \$1.00 for microfilm (images 1 inch high on standard 15 mm. motion picture film) or \$1.00 for photocopies (6 x 8 inches) readable without optical aid.

⁷ R. Flesch, "Measuring the Level of Abstraction," *Journal of Applied Psychology*, Vol. 34 (December, 1950), pp. 384-90, and D. W. Twedt, "A Table for Use with Flesch's Level of Abstraction Readability Formula," *Journal of Applied Psychology*, Vol. 35 (June, 1951), pp. 157-59.

TABLE 1

CORRELATIONS WITH READERSHIP, MEANS, AND STANDARD DEVIATIONS OF
34 ADVERTISING VARIABLES

<i>Variable</i>	<i>r</i>	<i>M</i>	<i>σ</i>
K. Criterion (Per Cent Readership)*	—	26.7	16.6
<i>Mechanical Variables</i>			
1. Number of pages or size of advertisement	.62†	.6	.6
2. Width-height ratio of advertisement	.15	.6	.3
3. Number of colors	.37†	2.3	.5
4. Number of separate illustrations	.28†	2.7	3.9
5. Square inches of illustration	.67†	12.4	17.5
6. Proportion of illustration	.52	24.3	16.4
7. Number of type styles	.06	4.9	1.6
8. Number of type sizes	.28†	5.6	1.6
9. Point size of largest type	.49†	38.9	23.1
10. Point size of headlines (wtd. avg.)	.43†	20.7	12.1
11. Largest type: product identification	.39†	27.7	19.2
12. Point size of main body copy	.24†	7.9	1.8
13. Pica width of copy measure (wtd. avg.)	.35†	14.6	5.5
14. Number of copy blocks	.30†	6.7	4.0
15. Layout deviation (±) from 90°	-.13‡	12.3	6.7
<i>Content Variables</i>			
16. Flesch readability scores	.00	46.9	10.7
17. Flesch abstraction level scores	.01	27.3	7.1
18. Number of words in advertisement	.31†	169.3	162.0
19. Number of words in headlines	.10	11.3	13.8
20. Number of product identifications	.40†	6.8	5.3
21. Number of product facts	.19†	16.5	11.5
22. Number of product benefits	.29†	11.6	8.9
23. Number of pictorial benefits	.18	1.6	2.2
24. Number of benefits in headlines	-.05	1.7	1.8
25. Number of benefits in body copy	.28	8.6	6.7
26. Number of pictures of product in use	.33†	2.3	1.7
27. Directions for getting more details	-.09	2.0	.7
28. News value ratings	.29	1.6	.7
29. Readership of surround	.32†§	30.5	12.6
30. Number of similar ads in issue	-.04	18.3	16.6
31. Previous schedule: $\frac{1}{50} + 1949$.47†	5.0	5.8
32. Previous schedule: $\frac{1}{50} + 1949 - 48$.45	9.2	13.3
33. Previous schedule: $\frac{1}{50} + 1949 - 48 - 47$.45	13.1	17.7
34. Brad-Vern totals	.23†	38.3	105.0

* *r*'s based on advertisements in Continuing Studies of Business Papers No. 2, *American Builder*, issue of February, 1950. *N* = 137 unless otherwise indicated.

† Significant *r*'s which are included in the 20 × 20 correlation matrix.

‡ Based on 57 full-page advertisements.

§ Based on 34 advertisements.

terion, 19 variables were selected on the basis of significant correlation with the criterion. In Table 1, variables 6, 23, 25, 32, and 33 were not included in the correlation matrix because they were not independent of other variables which were included. Variable 28 was not included

TABLE 2

CORRELATION MATRIX OF PRODUCT-MOMENT r 's FOR 20 ADVERTISING VARIABLES*

Variable	K	2	3	4	5	8	9	10	11	12	13	14	18	20	21	22	26	29	31
K Readership	62																		
2 Size of Ad	37	-07																	
3 Number of Colors	28	35	24																
4 Number Illustrations	67	71	21	25															
5 Sq. In. Illustration	28	54	21	22	27														
8 Number Type Sizes	49	64	32	21	45	54													
9 Largest Type	43	49	26	23	31	25	55												
10 Headline Size	39	41	28	04	33	41	66	26											
11 Largest Prod. Ident.	24	19	18	-07	17	25	23	25	25										
12 Body Type Size	35	25	28	-11	28	24	24	25	22	49									
13 Pica Width	30	57	20	54	23	40	29	22	14	-05	-01								
14 Number Copy Blocks	31	62	11	24	24	43	32	23	24	-03	12	61							
18 Number Words	40	61	18	33	36	-06	37	24	30	05	16	62	62						
20 Number Prod. Ident.	19	30	-02	25	26	16	19	-02	08	-03	-17	38	49	35					
21 Number Prod. Facts	29	51	11	30	22	37	31	16	20	-08	00	59	68	62	54				
22 Number Prod. Benefits	33	29	14	52	32	14	13	-02	08	-04	-13	45	16	29	35	37			
26 Pictures of Use	32	49	-11	12	38	20	44	45	15	24	22	-04	00	20	-38	-19	-17		
29 Surround	47	53	16	20	44	17	27	46	14	02	29	26	39	40	18	34	06	-18	
31 Previous Schedule	23	22	01	04	21	08	06	15	03	-04	04	09	19	17	02	22	10	-17	50
34 Brad-Vern Schedule																			

* Decimals omitted.

TABLE 3
FACTOR LOADINGS AND COMMUNALITIES OF 20 ADVERTISING VARIABLES*

Variable	Centroid Loadings						h^2_e	Rotated Loadings						h^2_r
	I	II	III	IV	V	VI		PC	S	T	In	F	A	
K Readership	74	-18	-21	-18	08	-12	6773	64	35	28	18	16	09	6766
5 Sq. In. Illustration	68	-13	-10	-13	14	-30	6158	51	48	25	06	23	04	6111
26 Pictures of Use	38	37	11	-50	05	-10	5559	51	23	-18	09	10	-44	5571
3 Number of Colors	33	-25	-17	-31	-15	13	3358	49	-07	23	11	-15	01	3326
1 Size of Ad	87	07	18	30	27	-10	9827	18	69	26	45	45	04	9827
21 Number Prod. Facts	35	48	15	-17	-21	-07	4533	20	37	-24	28	-12	-35	4498
34 Brad-Vern Schedules	26	19	-39	19	02	-27	3652	21	37	-25	-01	07	34	3641
9 Largest Type	71	-36	29	18	00	-06	7538	15	45	62	36	12	03	7543
11 Largest Prod. Ident.	51	-33	16	08	-14	-09	4287	18	34	48	21	-06	04	4277
8 Number Type Sizes	55	-18	46	11	-28	07	6419	01	37	47	47	-16	-19	6405
12 Body Type Size	26	-47	-09	-04	-02	04	3002	12	04	46	09	00	25	2982
10 Headline Size	55	-27	-10	18	06	04	4230	26	24	34	30	14	27	4233
18 Number Words	64	38	15	30	-18	17	7278	05	46	-10	71	00	-03	7291
14 Number Copy Blocks	62	43	21	-08	-04	27	6943	25	27	-13	65	10	-33	6937
20 Number Prod. Ident.	66	37	-11	14	04	25	6683	31	29	-20	64	19	06	6695
22 Number Prod. Benefits	60	59	13	14	-22	04	6965	14	52	-22	57	-06	-17	6958
29 Surround	24	-54	18	20	75	16	1.0097	-03	-13	61	10	76	17	1.0064
31 Previous Schedule	55	16	-53	25	04	-19	7092	42	46	-21	19	14	47	7087
13 Pica Width	34	-44	-33	08	-07	11	4415	35	-01	34	14	05	43	4452
4 Number Illust.	46	27	15	-34	18	07	4599	36	17	-05	27	25	-40	4564

* Decimals omitted. Boldface indicates factor on which each variable has its highest loading.

because of low reliability of judges. Means and standard deviations are included in Table 1 to provide the reader with some knowledge of the distributions from which the correlations were obtained. Product-moment intercorrelations were computed for these 19 variables plus the readership criterion, and incorporated in a 20 x 20 correlation matrix (Table 2). In general, the variables are positively correlated.

The correlation matrix was factor analyzed with Thurstone's complete centroid method.⁸ The resulting centroid matrix is shown in Table 3. Extraction was stopped with the sixth factor, since the product of the two highest loadings in factor VI ($.27 \times .30 = .08$) is only equal to the standard error of the original r between these two variables.

In order to give psychological meaning to these factor loadings, the arbitrary reference frame obtained by the centroid method was rotated by the graphical method. The resulting factors are orthogonal. Criteria of positive manifold and simple structure were observed wherever possible. Since most of the correlation coefficients are positive, it might be expected that a positive manifold could be obtained, and this actually was achieved with only a few exceptions.

INTERPRETATION OF FACTORS

In Table 3, boldface figures indicate the factor on which each test variable has its highest loading (a factor loading is the correlation between that test and that factor). Coefficients of determination (the squared factor loadings) give the percentage of variance of a given measurement which may be predicted by a particular factor. For example, the Readership variable has a loading of .64 on factor *PC*; thus $.64^2 = .41$, or 41 per cent of the variance in readership scores may be predicted from this single factor. Note that only two of the factors (Pictorial-Color and Size) have major loadings on Readership. Factor loadings below .20 are not usually considered significant; loadings between .30 and .40 may be important; if the projections are .40 or above, the loadings are considered significant.

Factor *PC* has high positive loadings on Readership (.64), Square inches of illustration (.51), Number of pictures showing the product in use (.51), Number of colors (.49), and Previous schedule of advertising (.42). The best measures of this factor are those involving Pictorial and Color aspects of advertisements, hence the factor designation *PC*.

Factor *S* has high loadings on Ad size (.69), Number of product

⁸L. L. Thurstone, *Multiple-Factor Analysis* (Chicago: University of Chicago Press, 1947), Chapter 8.

benefits (.52), Square inches of illustration (.48), Number of words (.46), Previous schedule of advertising (.46), and Largest type size (.45). Readership loading on Factor *S* is .35, or 12 per cent of the variance in readership scores is attributable to this factor, which seems to involve Size of advertisement.

Factor *T* has high loadings for Largest type size (.62), Readership of surround (.61), Largest type used for product identification (.48), Number of type sizes (.47), and Point size of main body copy (.46). In general, this factor seems to be associated with Typographic size and variety. Its Readership loading is .28, accounting for 8 per cent of readership variance.

Factor *In* has high loadings for Number of words (.71), Number of copy blocks (.65), Number of product identifications (.64), Number of product benefits (.57), Number of type sizes (.47), and Ad size (.45). This factor appears to be one of Information, and its loading on Readership is .18, accounting for only 3 per cent of readership variance.

Factor *F* has only two significant loadings: Readership of surround (.76) and Ad size (.45). The factor designation *F* is for Field—influence of the surrounding field or background against which the advertisement is seen. Another 3 per cent of readership variance is accounted for by this factor, which has a Readership loading of .16.

Factor *A* has significant loadings for Previous schedule (.47), Number of pictures of product in use (— .44), Pica width of copy measure (.43), and Number of illustrations (— .40). This factor is difficult to interpret, but tentatively it is called *A*, for Advertising schedule previously run. It accounts for less than 1 per cent of readership variance; the criterion loading is .09.

An important conclusion is that collectively these six factors account for two-thirds ($b^2r = .6766$) of the observed variance in readership scores of advertisements appearing in the February, 1950 issue of *American Builder*. *PC* alone accounts for 41 per cent of the variance in readership scores; *PC* and *S* together account for 53 per cent of the variance.

PREDICTION OF READERSHIP FROM MULTIPLE REGRESSION EQUATIONS

On the basis of the factor analysis, certain variables were chosen which seemed to be factorially purest, and which also offered most promise for prediction of advertising readership. Several combinations of these variables were tried in multiple regression equations, and the

following set of three (see Table 4) was selected as providing maximum prediction with minimum trouble of measurement.

$R_{1.234} = .77$ (where 1 = Predicted readership; 2 = Size of advertisement; 3 = Number of colors; 4 = Square inches of illustration). Correction for bias gives a shrunken R of .76. When nine variables

TABLE 4
CORRELATION MATRIX OF THREE ADVERTISING
VARIABLES

Variable	Size	Colors	Square Inches of Illustration
Size of advertisement		-.07	.71
Number of colors	-.07		.21
Square inches of illustration	.71	.21	

(numbers 1, 3, 11, 20, 21, 22, 29, 31, 34) were incorporated in a regression equation,⁹ $R = .79$. The gain of .03 is obviously not worth the time involved in making these additional measurements.

The best comparison of each variable's contribution to the variance in readership is found in column (4) of Table 5, where each beta weight is multiplied by the corresponding raw r . The regression co-

TABLE 5
CORRELATIONS WITH READERSHIP, β WEIGHTS, βr CROSS
PRODUCTS, AND REGRESSION COEFFICIENTS OF
THREE ADVERTISING VARIABLES

(1) Variable	(2) r_{1k}	(3) β	(4) r	(5) b_{1k}
Size of advertisement	.62	.441	.273	8.293
Number of colors	.37	.341	.126	3.869
Square inches of illustration	.67	.285	.191	.181

efficients, or optimal weights by which each variable must be multiplied to obtain a maximum multiple R , are given in column (5).

The regression formula computed from the *American Builder* data is:

$$X' = 10.456 + 8.293 (\text{Size of ad in pages}) + 3.869 (\text{Number of colors}) + .181 (\text{Square inches of illustration}),$$

⁹ R. L. Thorndike, *Personnel Selection* (New York: John Wiley and Sons, 1949), p. 340, gives an adaptation of the Kelley-Salisbury iterative solution for R which greatly facilitated these computations.

where $X' =$ predicted readership, and 10.456 is a correction for point of origin. Prediction of readership by this formula establishes relative differences rather than absolute readership scores, which are dependent upon the general readership level of a particular magazine.

The critical point of the study is now at hand: the factorial approach proved fruitful with the *American Builder* data, but what is the strength of the relationship between the mechanical variables of size, color, and amount of illustration and readership of advertising in other business magazines? Table 6 shows the product-moment correlation

TABLE 6
CORRELATIONS BETWEEN READERSHIP
SCORES OF ADVERTISEMENTS IN ARF
STUDIES, AND READERSHIP SCORES
PREDICTED FROM THE REGRESSION
FORMULA

Magazine Surveyed	<i>r</i>	Number of Advertisements
<i>Automotive Industries</i>	.58	131
<i>American Builder</i>	.76	137
<i>American Machinist</i>	.63	161
<i>Chemical Engineering</i>	.64	133
<i>Business Week</i>	.80	101
<i>Successful Farming</i> :		
Men readers	.77	217
Women readers	.73	217

coefficients between actual readership scores of advertisements in other ARF studies¹⁰ and readership scores predicted from the regression formula.

The mean r for the four business magazine studies is .66 (obtained by Fisher's z transformation). When *Business Week*¹¹ and *Successful Farming*¹² are included with the business magazine studies, the resulting mean r is .71.

DISCUSSION OF RESULTS

The three possible sources of variance in readership of advertisements are: (1) differences in the attention-getting power of the ad-

¹⁰ Advertising Research Foundation, *Continuing Study of Farm Publications*: No. 3, *Successful Farming*, issue of May, 1947; *Continuing Study of Business Papers*: Nos. 1, 2, 3, 4, *op. cit.*; *Study of Executive Management Publications*: No. 1, *Business Week*, issue of April 22, 1950.

¹¹ Advertising Research Foundation, *Study of Executive Management Publications*: No. 1, *op. cit.*

¹² Advertising Research Foundation, *Continuing Study of Farm Publications*: No. 3, *op. cit.*

vertisements, (2) differences in respondents' interests and purchasing readiness, and (3) chance errors in measurement.¹³ The present study is concerned only with readership variance attributable to differences in the advertisements, whether these differences are mechanical (size, color, illustration, etc.) or differences in content (number of facts, benefits, etc.).

When this analysis was begun, a possible outcome was that only a small part of the differences in readership might be accounted for by mechanical variables. A recent evaluation of the importance of content as against mechanical variables has been given by James D. Woolf,¹⁴ formerly vice-president of the J. Walter Thompson advertising agency, who stated:

It is my conviction that it isn't the size of the space that puts PULL into an advertisement. At least size is not the most vital consideration. Dr. Samuel Johnson said two centuries ago that "the soul of the advertisement is *the size of the promise.*" In other words, the size of the promised benefits.

I have no doubt that the huge units of space tend to achieve certain desirable results for the advertiser, they stimulate the sales force, they impress the trade, and very likely they enhance the importance of the product in the eyes of the reader. But they do not shape public opinion for a product when and if Dr. Johnson's "soul of the advertisement" is not in the copy. Huge size and red ink and thunderous pitch and clamor are not substitutes for promised benefits.

The italics are Woolf's. This position is clearly not supported by the findings of the present study. It should be remembered, however, that Woolf may refer to consumer advertising, and that he may exclude industrial advertising from consideration (although he does not make this distinction explicit in the article from which this quotation is taken), and secondly, his undefined "PULL" may refer to effect of the advertisement (which this study clearly does not attempt to predict) rather than to the size of its audience.

Lucas and Britt also stress the importance of content: ". . . the primary element in the success of all advertising copy is its *content* or substance. Most of the other factors are merely devices for making

¹³ This trichotomy is somewhat oversimplified; the possibility also exists that for two given advertisements, *A* might have greater immediate attention value than *B*, and yet *B* might be remembered more readily than *A* several days after *S*'s original exposure to *A* and *B*. Thus *A* might be said to have greater attention value, but *B* greater memorability. In the present study, these two variables (if they actually do exist independently) are confounded and their effects cannot be measured separately.

¹⁴ J. D. Woolf, "It Isn't Size That Puts Pull in Advertising," *Advertising Age*, Vol. 22 (April 30, 1951), p. 43.

the subject matter more visible, more palatable, and easier to comprehend."¹⁵

In an earlier experiment which was designed to measure the influence of mechanical variables on readership of newspaper advertisements, Ferguson concluded, "Contrary to popular and scientific belief it was found that there was no relationship between the size of an advertisement and its attention value."¹⁶ Ferguson's data are based on readership of a small daily newspaper, and again there may be real differences between readership of industrial advertisements in business magazines, and consumer advertisements in small daily newspapers.

Although the present conclusions as to the importance of the mechanical variables of size, color, and illustration are based primarily upon industrial advertising in business magazines, it is suggestive that the highest relationship between readership scores, and readership as predicted by the regression formula, was for *Business Week*, an executive management publication which is somewhere in between the business magazine edited for a particular industry or occupation, and the general magazine with almost universal appeal. Unless this r of .80 (Table 6) represents only a vagary of sampling, it is reasonable to assume that the regression weights given in Table 5, column (5) may prove useful in predicting the relative readership of advertisements in general magazines. The *Successful Farming* study, with r 's in the .70's (Table 6), also supports this assumption.

SUMMARY

Thirty-four advertising variables were defined, measured, and correlated with readership scores for 137 advertisements in the February, 1950 issue of the *American Builder*, a business magazine published primarily for building contractors. Criterion scores were obtained from the Advertising Research Foundation's Continuing Studies of business magazine readership.

Of these 34 variables, 19 were selected as most significantly correlated with the criterion. Product-moment intercorrelations were computed for these 19 variables plus the readership scores, and the resulting 20 x 20 correlation matrix was factor analyzed.

Six factors were found to be sufficient to account for the intercorrelations. Of these six factors, only two, PC (Pictorial-Color) and S (Size)

¹⁵ D. B. Lucas and S. H. Britt, *Advertising Psychology and Research* (New York: McGraw-Hill Book Co., 1950).

¹⁶ L. W. Ferguson, "The Importance of the Mechanical Features of an Advertisement," *Journal of Applied Psychology*, Vol. 19 (October, 1935), pp. 521-26.

have major loadings on Readership. The other factors are *T* (Typographic size and variety), *In* (Informational), *F* (Field factor, or the influence of the surrounding field of the advertisement), and *A* (Advertising schedule previously run). Collectively, these six factors account for two-thirds of the observed variance in readership scores of the advertisements. The *PC* and *S* factors alone account for 53 per cent of the variance in readership.

On the basis of the factor analysis, certain variables were chosen which seemed factorially purest, and a multiple regression equation was developed to predict readership of advertisements in other business magazines. A multiple *R* of .77 was obtained between readership and the following group of variables: size of advertisement, number of colors, and square inches of illustration.

The regression equation was employed to predict readership of advertisements in six other Advertising Research Foundation studies. Predicted readership scores were correlated with actual readership scores, and these validity coefficients ranged from .58 to .80, with an average *r* of .71.

29. WHAT MAKES A BEST SELLER?*

Why people are attracted to some books and not to others is a problem that can be approached in at least two ways—through an analysis of the content of the books themselves or through personal interviews with readers (and nonreaders) of different types of books. The former approach, "content analysis," is followed in this interesting article, which summarizes previous studies of this type and describes the application of the method in considerable detail.

Content analysis is a method which has received considerable use to provide clues to "why" questions in public opinion work, but has not been accorded too much attention in marketing research. One virtue of the method, as exemplified by its application here, is its ability to consider large numbers of variables from which a few relevant, or "most pertinent" ones can be selected. Quantification and use of the results for prediction then follows by some more precise method, in this case, discriminant analysis.

* Adapted from an article by John Harvey, Parsons College, "The Content Characteristics of Best-Selling Novels," *Public Opinion Quarterly*, Vol. 17 (Spring, 1953), pp. 91-114.

"What makes a best seller?" is a perennial question to which countless answers have been offered by publishers, booksellers, librarians, and authors. None of their hypotheses and opinions, however, have demonstrated why one book sells better than another, nor is it yet possible to predict the sales of any book on the basis of tested criteria.

The study on which this article is based confined itself to best-selling novels.¹ The popularity of these novels might be explained in terms of their timeliness, their authors, their sales promotion, or in the prestige attached to having read them. By contrast, the present study concentrated on the content of novels in an attempt to isolate content characteristics differentiating best sellers from similar novels which failed to sell so well.

PREVIOUS APPROACHES TO THE PROBLEM

There have been numerous approaches to the "why" of best sellers. The historical approach is best illustrated by Frank Luther Mott's history of American best sellers, which contains two chapters on causation.² Mott found sensationalism, themes of religion and adventure, and a sentimental treatment of subject matter recurring frequently in best sellers. His major thesis, however, was the frequently encountered one that there is no "best seller formula" and consequently that lists of important variables are of little value.

George Stevens, of J. B. Lippincott Company, sought to explain sales in terms of publicity advantages.³ He detailed figures on advertising appropriations, discussed the promotional techniques used with several well-known best sellers, and mentioned the importance of such advantages as publicity off the book page, condensation in the *Reader's Digest*, distribution by a book club, and a controversial theme.

Emmanuel Haldeman-Julius, with a somewhat different approach because of differing marketing problems, found that many of his "Little Blue Books" sold much better when given new titles.⁴ His retitling played up themes of love, sex, religion, self-improvement, and humor.

Literary critics have made numerous contributions to the best seller literature. As examples, we might mention the articles by Granville

¹ J. F. Harvey, *The Content Characteristics of Best-Selling Novels* (Ph.D. thesis, University of Chicago, 1949).

² F. L. Mott, *Golden Multitudes; The Story of Best Sellers in the United States* (New York: Macmillan, 1947).

³ G. Stevens, *Lincoln's Doctor's Dog, and Other Famous Best Sellers* (Philadelphia: J. B. Lippincott Company, 1938).

⁴ E. Haldeman-Julius, *The First Hundred Million* (New York: Simon and Schuster, 1928), pp. 138-78.

Hicks and Edward Weeks.⁵ Hicks read forty novels from the period 1930 through 1934 before arriving at his conclusions. The formula he recommended included ". . . a lively story, largely romantic in theme and setting with conventional characters and plot and some pretention to a message or thesis, apparently profound but really commonplace."⁶

Weeks, instead of studying a group of novels, studied only *Gone with the Wind*. He estimated its success to be due to the following factors in the proportions indicated: timeliness, 45 per cent; emotion, 25 per cent; characterization, 15 per cent; invention, 10 per cent; and advertising, 5 per cent. His interpretation of the novel's appeal emphasized the vitality of the major characters, the necessity of including an impressively large number of minor characters, the use of superlatives, and the use of overlong scenes to hold the reader's interest. His one-sentence formula ran as follows: ". . . the novelist captures the ideas in the air at the time and puts them into words . . ."⁷

The limitations of the above studies are readily apparent. They either failed to examine the books with sufficient thoroughness, conducted analyses which were too subjective, or else examined too few books for any sensible generalizations to be made. The studies by Joseph Kappel and J. V. M. Berreman are more important because these pitfalls were avoided and reliable analyses obtained. Joseph Kappel studied the literary quality of the best sellers and book club selections of the last twenty years.⁸ He used *Book Review Digest* plus and minus ratings of reviews as his measure of literary quality and obtained from them annual index numbers showing the degree of favorability shown by the reviews. These index numbers were obtained for all Book of the Month Club and Literary Guild selections, for a random sample of each year's new books, and for all best sellers (both fiction and nonfiction) of the years 1926 through 1941. The annual favorability index numbers were then plotted on charts to show comparisons among the groups of books for each year. These charts showed that, from a literary standpoint, best sellers as well as Book of the Month Club and Literary Guild selections ranked above the random sample (or average) each year. Kappel's study thus tended to suggest that neither the best

⁵ G. Hicks, "The Mystery of the Best Seller," *English Journal*, Vol. 23 (October, 1934), pp. 621-29; E. Weeks, "What Makes a Book a Best Seller?" *New York Times Book Review*, Vol. 41 (December 20, 1936), pp. 2, 15.

⁶ Hicks, *op. cit.*, p. 626.

⁷ Weeks, *op. cit.*, p. 2.

⁸ J. W. Kappel, "Book Clubs and the Evaluation of Books," *Public Opinion Quarterly*, Vol. 12 (Summer, 1948), pp. 243-52.

sellers nor the book clubs have reduced the general quality of books in recent years.

In a study entitled *Factors Affecting the Sale of Modern Books of Fiction*, J. V. M. Berreman examined the role of both content and noncontent.⁹ This study merits careful consideration because of its thoroughness and objectivity, as well as for its close relation to the present study. Berreman's purpose was the prediction of book sales. He examined 234 best- and poor-selling novels published between 1933 and 1938. Their sales were measured by the frequency of their appearances in the weekly best seller charts of the *New York Herald*

TABLE 1
BERREMAN'S CORRELATIONS BETWEEN PUBLICITY
FACTORS AND SALE

Publicity Factors	1935 Random Sample for Publisher's Lists		1933-38 Best Seller Titles	
	N	r	N	r
Advertising	77	+.79	50	+.28
Author prestige	228	.65	71	.48
Review wordage	77	.52	50	.10
Review favorability	77	.12	50	.09

Tribune book review section.¹⁰ Berreman found close relationships between several of the noncontent (or publicity) factors and sales, although no causal relationships were established. These relationships were sufficiently close, however, to allow him to combine all of the important factors into a formula which yielded a correlation of $+.77$ with the sale of a selected group of novels.

Table 1 shows Berreman's correlations to be high only when he was distinguishing between best and poor sellers, as in his 1935 group. He could successfully predict a best or a poor seller from known differences in its publicity factors, but he could not equally predict differences between the sales of the top best sellers (the "super" best sellers) and those lower on the best seller lists. In short, his publicity factors explained only gross differences in sales.

Berreman devoted but little attention to content differences. On the basis of reviews, he grouped sixty novels into twelve classes by setting, theme, and treatment of theme. He concluded that content probably affected the sale of poor sellers which were well marketed as well as

⁹ J. V. M. Berreman, *Factors Affecting the Sale of Modern Books of Fiction; A Study of Social Psychology* (Ph.D. thesis, Stanford University, 1940).

¹⁰ The validity of the measure of sale was tested successfully.

best sellers which sold well despite feeble efforts to market them. He also pointed out the probable usefulness of content in any attempt to improve the correlations with sale, shown in Table 1.

The present study is essentially an extension and supplement to the Berreman study. It begins where Berreman stopped; namely, with the analysis of differences in the content of novels and with the relating of such differences to their sales. Berreman's findings were accepted for what they were worth so that the present study arbitrarily excluded publicity factors from consideration and concentrated its attention on content.

THE VARIABLES

The core of the present study was a comparison of books in terms of certain variables. The purpose of the analysis was to discover recurring differences between the scores of best and poor sellers on each variable and on each group of variables. The variables used for the analysis are best described in terms of the six main groups into which they naturally fall. The groups and the numbers of variables included are as follows:

Action	15 different variables
Emotion	50 different variables
Personalities of the major characters	750 different variables
Plot themes	350 different variables
Romanticization	50 different variables
Simplicity	32 different variables

Action. Action referred to any change or movement of the plot, setting, or characters. An action was shown whenever the plot themes changed, part of the setting moved (as for instance a door slammed), the scene of action changed, or a character moved. Such measures as the following were used: frequency of movement by each major character per 100 character lines, the frequency with which the reader's attention was shifted from character to character, and the frequency per page with which the scene shifted to a new location.

Emotion. The pattern of emotion referred to the amount, kind, and use of the emotion expressed in each novel. It included the amount of certain basic emotions shown by each character, the frequency of change from one emotion to another, and the percentage of the lines with an emotional charge.

Personalities of the Major Characters. The "personality" group of variables included many different aspects of each character's life: for example, the character's age, occupation, education, nationality, social status, social traits, morals, and prominence in the novel.

Plot Themes. The pattern of the plot themes referred to the inevitably differing themes used in different novels. The list of themes was taken from the table of contents of Elbert Lenrow's *Readers Guide to Prose Fiction*, which contains approximately 350 different themes.¹¹ The coverage of this list is indicated by the following outline: entertainment and escape, the individual and his personal environment (people and their personal problems), and the individual and his social environment (social, political, economic, vocational, religious, and philosophical problems).

Romanticization. Examples of variables in the romanticization category included extremes of each character's wealth or poverty, the contrast between income level at the beginning and end of the novel, the physical attractiveness of each character, his moral perfection, the luxury or poverty of the setting, and fortuitous events (such as receiving a million dollars without warning). It was the combination of unusually high or low scores for several of these variables that revealed the degree of romanticization.

Simplicity. The variables measuring simplicity tested the ease with which the average reader could move through the novel. These variables were related either to: (1) the novel's style, (2) its plot, (3) its major characters, or (4) its setting. Simplicity of style was measured by such variables as the total wordage of the novel, the average chapter length, the frequency of abstract idea words, the number of lines and paragraphs per 100 words, and the readability of the novel. The principal measure of plot simplicity was the number of Lenrow themes in each novel.

Variables relating to the major characters included the number of such characters in the novel, the number of goals for each character, the derivation of their names, and the allotment of lines between setting and characters. The five most important categories of major characters were the central male character and his best friend, the central female character and her best friend, and the villain.

The variables related to the setting included the degree of luxury, the number of geographical locations per 1,000 words, and the frequency of changes of setting.

Sources of the Variables. The variables were taken from a variety of sources. However, three kinds of sources predominated: subjective analyses of best sellers by literary critics (such as the Weeks, Hicks, and

¹¹ E. Lenrow, *Readers Guide to Prose Fiction* (New York: Appleton-Century, 1940), pp. vii-xi.

Bennett¹² studies); general outlines for analyzing fiction (such as the Krieg¹³ and Muller¹⁴ outlines and the Lenrow classification of themes); and, finally, my own observations as the analysis progressed.

The literary critics were a somewhat barren source. Nevertheless, Katherine Fullerton Gerould (in criticizing the leading female novelists) pointed out the importance to the female reader of detailed descriptions of social rituals, such as dinners, introductions, parties, and dressing scenes;¹⁵ Edward Weeks contributed such variables as the frequency of shifts in character and setting and the importance of numerous minor characters;¹⁶ George Orwell, in discussing popular boys' magazines, stressed the importance of action;¹⁷ and Granville Hicks contributed an important theme—the moralizing theme—and suggested the importance of the novel's length.¹⁸

Most of the variables were supplied by the second source—the Krieg, Muller, and Lenrow outlines. The Krieg outline contained many variables used in analyzing the major characters—for example, the characters' social and character traits, their attitudes toward themselves, toward others, and toward social institutions. Muller suggested sociological variables relating to the major characters, such as social and economic status, ethicality, social perspective, age, and occupation. The use of Lenrow's list of themes has already been described.

There were certain problems to be solved before the first two groups of sources could be used satisfactorily. The chief problem with the literary critic's contributions was that of selecting variables which could be clearly defined and reliably measured. Since the Krieg and Muller outlines, on the other hand, were developed for other purposes, the chief concern was to select variables which would fit the present problem.

Finally, the variables originated by the present study were derived directly from the novels' content—whether, for example, the novel was written in the first or third person; relative amounts of space de-

¹² E. A. Bennett, *Fame and Fiction: An Enquiry into Certain Popularities* (London: G. Richards, 1901).

¹³ L. L. Krieg, *A Suggested Method of Analysing Children's Fiction Reading* (M.A. thesis, University of Chicago, 1943), pp. 13–32.

¹⁴ H. Muller, *Social Stratification in Magazine Fiction and Its Relation to the Socio-Economic Status of Readers* (Ph.D. thesis, University of Chicago, 1942), pp. 59–65.

¹⁵ K. F. Gerould, "Feminine Fiction," *Saturday Review of Literature*, Vol. 13 (April 11, 1936), pp. 3–4, 15.

¹⁶ Weeks, *op. cit.*, p. 2.

¹⁷ G. Orwell, *Dickens, Dali, and Others; Studies in Popular Culture* (New York: Reynal and Hitchcock, 1946), pp. 100–101.

¹⁸ Hicks, *op. cit.*, pp. 626–29.

voted to author description, character thought, and character conversation; the recency of events; lines of emotion for each major character; and the frequency of profanity.

It will be recognized that the sources were used in any way that would benefit the study. Variables were lifted out of context whenever they could be made to serve the present purposes. The final list contained as many variables as could be supported by plausible hypotheses and analyzed objectively.

THE SAMPLING PROBLEM

Sampling for Certain Variables. Some of the variables could be measured feasibly only by sampling each novel; trying to measure them for the entire novel would have required a prohibitive expenditure of time and labor. Readability, for instance, required measurement only on a small proportion of the novel's sentences. Consequently, some 510 of the variables were measured by sampling, while the remainder were measured for the novel as a whole.

Thirty pages of each novel were sampled with the page numbers selected from the Tippett tables of random sampling numbers.¹⁹ The Tippett tables were developed to provide a list of 10,000 numbers which had been selected by a completely random method. The numbers in the tables were followed in order until thirty were picked up which were within the page limits of the novel being analyzed. The pages corresponding to these numbers were then analyzed for the variables in question.

Reliability of the Sampling Method. Two methods were used to check the reliability of the sampling. The first involved an attempt to prove statistically whether or not all the variables were adequately sampled in thirty pages, while the second involved the use of new thirty-page samples from several novels. The attempt to prove whether or not all variables could be adequately sampled on thirty pages of a novel yielded results indicating only that a random sample of 400 words should be adequate.

It would obviously be impossible to rate more than a few of the variables on the basis of a random sample of 400 isolated words, because there would be no continuity, no story to follow. The alternative of analyzing every word on a random sample of thirty pages would seem satisfactory, although this method would require more than 400

¹⁹ L. H. C. Tippett, *Random Sampling Numbers*, Department of Applied Statistics (Computing Section), University of London, University College. Tracts for Computers . . . no. 15 (London: Cambridge University Press, 1927).

words to sample the novel as well as could be done with 400 randomly selected words. Since it was impossible to say how many more words would be needed, it was necessary to assume that thirty randomly selected pages, usually containing about 10,000 words, was enough.

In the second test, another randomly selected thirty-page sample was drawn from the Tippet tables for each of five novels. No pages were duplicated between the two samples. Table 2 shows the per cent

TABLE 2
RELIABILITY OF THE METHOD OF SAMPLING THE NOVELS

<i>Factors</i>	<i>Mann: Joseph the Provider</i>	<i>Marshall: Duchess Hotspur</i>	<i>Remarque: All Quiet on the Western Front</i>	<i>Stribling: The Store</i>	<i>Warren: All the King's Men</i>
1. Lines of central male character's level 1 emotion					
a. Unhappiness	97%	96%	92%	96%	97%
b. Love	100	100	100	99	100
c. Anger	98	99	100	97	96
2. Number of major characters toward whom central male character's attitude is					
a. Affectionate	100	100	100	100	89
b. Respectful	100	89	100	100	89
3. Sentimental theme	90	93	97	90	93

Explanation: This table shows the results of rating 30 new sample pages from each of five novels for six of the variables found to be good discriminators. The percentages show the proportion of lines, pages, or characters which were rated the same way, or produced the same ratings, for the new sample as for the original sample. In other words, the table shows the percentage of agreement between the new and original samples for each of the six factors. Totals of 4850 lines and 29 characters were rated on 150 new sample pages.

of correct out of total ratings to be satisfactorily high for all six variables. This means that the six variables occurred with the same frequency in the new sample as in the original sample, thereby encouraging confidence in the reliability of the original sample.

Reliability of the Variables. The objectivity of the present description of content was also proven for the above six variables in tests of their reliability. The tests indicated whether these variables were being measured objectively or subjectively—whether they could be measured consistently by the same and other raters or whether they were ill-defined and the measurement was erratic.

The reliability of the ratings could be tested by two methods: by

examining the consistency of the original rater, and by comparing the observations of several raters. Both methods were employed in the present tests. The test of internal consistency—testing the percentage of lines classified the same way several months later by the original rater—indicated a fairly high degree of agreement between the two ratings.

For the test of consistency among several raters, samples of the six variables were selected by the original rater on seven pages in one of the novels analyzed early in the study, Bottome's *Mortal Storm*. Definitions were then explained to three additional raters and their ratings obtained for the seven pages. All six variables showed a high degree of reliability.

Pearson's Chi-Square (χ^2) test of the dispersion of individual scores about an expected score was also applied to the observations of the four raters. This test measured the likelihood that the differences between raters in the scores for each variable were actually significant. If the scores were not significantly different, then the variations were due to chance factors alone.

The mean score for all four raters was used as the "expected" score for each variable. The mean was used for this purpose because there was no expected score in the sense that one rating or one rater was necessarily correct and all others incorrect. For each of the six variables tested, results indicated that there was a probability of from .60 to 1.00 that the differences shown between the raters were due to chance. When the variables were lumped together and their total tested, results were equally favorable and the ratings reliable. There was a probability of .80 to .90 that the differences were due only to chance factors; these variables, therefore, were rated reliably in this test.

THE NOVELS

A best-selling novel was defined as any novel which ranked among the top ten novels in sale for a year, as reported in Alice Hackett's book, *Fifty Years of Best Sellers, 1895-1945*,²⁰ or in her annual summary articles in *Publishers Weekly*. A poor-selling novel was one which did not appear on these lists.

Sampling the Best and Poor Seller Universes. The time required for analyzing each novel (100 to 150 hours at the beginning of the analysis) made prohibitive the use of all eligible best and poor

²⁰ A. P. Hackett, *Fifty Years of Best Sellers, 1895-1945* (New York: R. R. Bowker Co., 1945).

sellers; some method was required for sampling the best and poor seller universes. Berreman's findings indicated that a random sample would not be desirable for this purpose; he showed that sales differences were positively correlated with differences in the amounts of money, ingenuity, and other resources used to promote books and publicize them. Hence, a random sample of novels would not allow identification of the content elements influencing sale unless publicity factors were eliminated, neutralized, or held constant in the sampling operation. It was therefore required that the poor sellers used in the sample be as well endowed with publicity advantages as were the best sellers; in other words, best and poor sellers were equated or matched on publicity factors so that the influence of content on sale could be studied directly. In addition to the influence of publicity factors, it seemed probable that such factors as the novels' themes, wordages, publication dates, and perhaps even favorability of reviews were also involved. Consequently, these factors were matched for best and poor sellers in the same way that publicity factors were matched.

This matching of novels constituted a "matched sample" from the best and poor seller universes. Use of the matched sample suggested the desirability of matching and analyzing best and poor sellers, title by title. Accordingly, each best seller was closely matched with a poor seller to form a pair for purposes of analysis.

Selecting the Matched Sample: Publicity Factors. The closest feasible matching of publicity factors required use of the weights which Berreman developed for his formula. Berreman did not explain his weights in detail, nor how he arrived at his values, nor did he define his categories. It was therefore necessary to adapt his procedure (as it could be understood) to the present study, and in some cases even to modify the original weights. Each novel was scored on each publicity factor separately and a weight assigned; then all the weights were summed for the novel and the total taken as the novel's final score on publicity. A criterion was adopted for each pair of novels which stated that the total of the poor seller scores on publicity must always equal, if not exceed, that of the best seller. Pairs in which this criterion was not met were eliminated before the analysis started. The publicity factors, sources of information on them, and the weighting used in the present study are given in Table 3; total possible score is 9.5 points.

Matching Novels by Major Theme. To obtain closely matched samples of best- and poor-selling novels, it was also necessary to control the major theme; matching two novels with different major themes would have defeated the present plan of comparison and in-

TABLE 3
SCORE CARD FOR WEIGHTING ON PUBLICITY FACTORS

Factor	Source	Weighting
A. Book club selection:		
Book-of-the-Month Club	<i>Publishers Weekly</i>	
1926-1930		1.0
1931-1946		3.0
Literary Guild		
1927-1936		0.0
1937-1943		1.0
1944-1946		3.0
B. Pulitzer Prize	<i>World Almanac</i>	2.0
C. Prepublication and first month advertising in the following journals:	Direct measurement	
<i>New York Herald Tribune</i> Sunday book review section		
<i>New York Times</i> Sunday book review section		
<i>Publishers Weekly</i>		
<i>Saturday Review of Literature</i>		
1. Total of more than eight pages		1.5
2. Total of 5 to 7.9 pages		1.0
3. Total of 3 to 4.9 pages		0.5
D. Author popularity: Inclusion of author on Hackett lists during the preceding ten years for best selling novels with similar themes	Hackett: <i>Fifty Years of Best Sellers</i>	1.0
E. Wordage of reviews in the following journals:	Direct measurement	
<i>Nation</i>		
<i>New York Herald Tribune</i> Sunday book review section		
<i>New York Times</i> Sunday book review section		
<i>New Yorker</i>		
<i>Saturday Review of Literature</i>		
<i>Springfield Republican</i>		
<i>Time</i>		
1. Total of more than 5500 words		1.0
2. Total of 3500 to 5499 words		0.5
F. Serialization or condensation in a national periodical	<i>Fiction Catalog</i> and <i>Readers Guide</i>	0.5
G. Other publicity factors aiding sale:		
1. Author reputation in some other field, or	<i>Twentieth Century Authors</i>	
2. Total of ten or more references, exclusive of reviews, in <i>Readers Guide</i> and <i>New York Times Index</i>	Direct measurement	0.5

vited the criticism that the novels were not written with the same purpose or to appeal to the same audience. On the other hand, it was recognized that major theme might be important in influencing sale and that it was unfortunate to lose this possibly good discriminator. But it was assumed to be more desirable to evaluate the many other characteristics with major theme controlled than to restrict the com-

parisons to major theme alone. Accordingly, the rule was adopted that the novels of each pair have similar major themes, similar geographic and historic settings, and similar major characters.

Limitations on Publication Dates. In order to insure a degree of homogeneity among the novels and to increase uniformity of taste among the readers, time limitations on publication dates were adopted, both for the novels as a whole and for the novels in each pair. The limitation adopted for the entire group of novels required that they be published between 1930 and 1946, inclusive, while the time span between the novels of each pair was limited to five years.

Additional Factors. Three other factors were considered in the matching procedure: wordage of the novel, order of publication of the novels in each pair, and favorability of reviews. The first two factors were found to have a positive relation to sale. There was some advantage associated with a large wordage and with being published prior to the other novel in a pair. Berreman did not report these factors, however, and their importance was not discovered until after the matching procedure was under way, so it was possible to control them for only the first eight pairs. These were so selected that neither best nor poor seller was heavily favored on either factor. To have carried the matching beyond these first eight pairs would have seriously reduced the number of pairs remaining, so no further control was possible. The third factor, review favorability, was checked for all pairs of novels, and the best and poor sellers were found to be equally benefited by it.

Sales Ratio in Each Pair. The final factor to be controlled was the sales levels of the novels in each pair. The foregoing definition of best and poor sellers merely specified that the best seller be on the annual Hackett lists of the top ten novels and that the poor seller not be on them. This would allow matching a best seller in tenth place with a poor seller in eleventh place, just off the list. Such a pairing would scarcely serve to compare a best with a poor seller because their sales would have been so nearly the same. It was accordingly decided that the best and poor seller in each pair should be separated by at least a 4:1 ratio in sale.

In the absence of any reliable figures on actual sale, an approximation was obtained by summing the monthly decimal scores for each novel in *Publishers Weekly* and converting them to whole numbers; the total thus obtained (hereafter referred to as the *PW* score) served as an estimate of the sale of each novel in relation to other novels. These monthly decimals or percentages were collected for the duration of each novel's stay on the lists, but not exceeding two years following

TABLE 4
 SCORES OF LANGLEY'S *A Lion Is in the Streets*
 AND WARREN'S *All the King's Men*
 ON THE MATCHING FACTORS

<i>Langley (the Best Seller)</i>	<i>Factors</i>	<i>Warren (the Poor Seller)</i>
	<i>Publicity Factors</i>	
Not a selection	Book club selection	Not a selection
No	Pulitzer prize	Won prize for 1946
	Advertising pages:	
4 pages	<i>Publishers Weekly</i>	2 pages
1 ½ pages	<i>New York Herald Tribune</i>	¾ pages
1 page	<i>New York Times</i>	1 page
1 page	<i>Saturday Review of Literature</i>	1 page
Total pages: 7 ½		Total pages: 4 ¾
Never	Author prestige (on previous Hackett lists)	Never
	Review wordage:	
0 words	<i>Nation</i>	1400 words
1200 words	<i>New York Herald Tribune</i>	1900 words
1050 words	<i>New York Times</i>	1150 words
200 words	<i>New Yorker</i>	100 words
950 words	<i>Saturday Review of Literature</i>	850 words
500 words	<i>Springfield Republican</i>	800 words
0 words	<i>Time</i>	750 words
Total words: 3900		Total words: 6950
Never	Serialization	Never
Never	Condensation	Never
None	Author reputation in another field	None
	References to novel:	
0 references	<i>New York Times Index</i>	5 references
3 references	<i>Readers Guide</i>	8 references
Total references: 3		Total references: 13
	<i>Non-Publicity Factors</i>	
6th in 1945	Hackett list	On no lists
Huey Long biography	Major theme	Huey Long biography
May, 1945	From 1930-1946 period	August, 1946
15 month interval	Five-year interval
150,000 words	Wordage in novel	230,000 words
Published in 1945	Order of publication	Published in 1946
280 <i>PW</i> score	4:1 ratio in <i>PW</i> score	20 <i>PW</i> score
	Review favorability:	
5 reviews	+	7 reviews
3 reviews	+-	3 reviews
0 reviews	-+	0 reviews
0 reviews	-	1 review

its publication date; the two-year limit gave every novel the same chance to make its score.

Examples of the Pairing and Analysis Procedures. An example of this scoring for a pair of novels would be in order here. One pair was Langley's *A Lion Is in the Streets* (the best seller), and Warren's

All the King's Men (the poor seller). Langley's book appeared on the *Publishers Weekly* charts for six months, from June through November of 1945. During this period its monthly percentages were 43, 65, 65, 44, 37, and 26. When these scores are summed, the total, 280, is the final *PW* score for this novel. Warren's novel appeared on the charts for only one month, September, 1946; its score for this month was 20 per cent. Therefore, the best seller scoring 280 and the poor seller scoring 20 make a ratio of 14:1 which is sufficiently high to meet the criterion.

To clarify the pairing procedure, Table 4 shows the raw scores of the two novels on all of the matching factors. Comparison of the raw scores in Table 4 to the weights in Table 3 will allow calculation of the scores on the publicity factors. Langley scored as follows: advertising, 1.0, and reviewing, 0.5; while Warren scored 2.0 on the Pulitzer prize, 0.5 on advertising, 1.0 on reviewing, and 0.5 on references. These totals are then 1.5 for Langley and 4.0 for Warren, thereby meeting the criterion that the poor seller score equal if not exceed that of the best seller. In general, the Langley-Warren pair was well matched on all factors.

The Analysis Procedure. The analysis procedure consisted principally of coding the variables on each sample page, totaling and transferring them to coding sheets, and transcribing the final scores (usually expressed as frequencies per 100 words or a percentage) to the result sheets. It was by comparing the final scores of each pair on each variable on the result sheets and then applying the 50 per cent criterion that the retention or elimination of the variable was judged at frequent intervals during the analysis.²¹

FINDINGS: THE IMPORTANT VARIABLES

Table 5 presents the major findings of the study and shows that only sixteen variables survived all eliminations. These variables were therefore the best single-variable predictors of sale. On the basis of these findings, the following generalizations seem warranted for the novels in this study:

1. Emotion is a major ingredient of the best seller. This generalization is warranted because the best seller shows:
 - a. More sentimentality (variable 4).

²¹ The 50 per cent criterion stated that, to be retained for further analysis, a variable should have scored significantly higher or lower on the best seller in at least half of the pairs already analyzed.

TABLE 5
VARIABLES STILL IMPORTANT AT THE END OF THE ANALYSIS

Code Number and Variable	Final Score*		
	+	0	-
<i>Style</i>			
1. Readability (Flesch)	1	12	9
2. Average chapter wordage	11	5	6
3. Recency of events in the novel	8	4	10
4. Sentimental theme	9	8	5
5. Moralizing theme	7	15	0
<i>The Central Male Character</i>			
6. Sensationalism as measured by exclamation points	11	2	9
7. Affectionate attitude toward other characters	9	10	3
8. Respectful attitude toward other characters	6	3	13
9. Total of level 1 emotion†	12	4	6
10. Total unhappiness (including both levels of intensity)	12	3	7
11. Unhappiness at level 1 (less intense)	11	3	8
12. Total anger (including both levels of intensity)	11	4	7
13. Anger at level 2 (more intense)	6	15	1
14. Anger at level 1 (less intense)	9	6	7
<i>Noncontent Variables‡</i>			
15. Prior publication date	18	0	6
16. Wordage of the novel	9	9	4

* These scores are to be interpreted as follows: a score of 1-12-9 means that in one pair of novels the best seller's score was significantly larger than that of the poor seller, in twelve pairs the scores were even, and in nine pairs the poor seller's score was significantly larger than that of the best seller.

† Specific emotions were scored in five different ways: level 2 indicated very strong emotion with the character in an extreme emotional state; level 1 indicated emotion of some strength but not at an extremely high level; total emotion of each kind (unhappiness, love, anger, etc.) contained a summation of the amount of emotion shown in both levels 1 and 2; all emotion at level 1 and all emotion at level 2 were summed; and finally the totals for level 1 and level 2 were summed to obtain a score for all emotion displayed in the 30 page sample of each novel.

‡ Variables 15 and 16 were segregated from the other variables because they were only partially related to content; they contained both content and noncontent elements and therefore did not meet the rigorous interpretation of a content variable used in this study.

Regarding prior publication, it was hypothesized that, of two novels on the same theme, the first to be published probably "soaked up" all of the market for that particular theme or treatment of it and left the second novel with little or no market at all. This involved the content of the novel in that it involved its theme, but publication date itself had nothing to do with content.

Wordage involved the amount of the content in a novel, but noncontent factors were present because a large wordage apparently encouraged readers to buy with the idea that they were getting more for their money, without regard to the ingredients of the novel. The readers seemed to believe that their pleasure would last longer with a long novel than with a short one.

- b. More central male character sensationalism (variable 6).
- c. More central male character affection toward other characters (variable 7).
- d. More central male character level 1 emotion (variable 9).
- e. More of the specific emotions listed in variables 10 through 14.

2. Certain characteristics of the central male character (variables 6 through 14) are either positively or negatively associated with sale.
3. There are certain stylistic and thematic differences between best and poor sellers (variables 1 through 5).

Retention of Variables Failing to Meet the 50 per cent Criterion. After twenty-two pairs of novels had been analyzed, the 50 per cent criterion was a plus or minus score of eleven. This meant that a variable's scores had to show a significant plus or a significant minus difference between best and poor seller in eleven of the twenty-two pairs. Table 5 shows that seven of the fourteen acceptable variables did not meet this criterion. Why were they retained? The reason for retention was primarily one of anticipated value: It was anticipated

TABLE 6
STATUS OF THE GROUPS OF VARIABLES AT THE BEGINNING AND
END OF THE ANALYSIS*

<i>Groups</i>	<i>Number of Variables at Beginning of Analysis</i>	<i>Number of Variables at End of Analysis</i>
Action	15	0
Emotion	54	10
Personalities	75	9
Plot themes	350	2
Romanticization	50	0
Simplicity	32	1

* Since several of the variables belonged to more than one group, the sums of the two columns equal more than the number of different variables included.

that they might be good discriminators when used in the formulas. Each was retained for at least one of three reasons: that the variable's scores showed a strong tendency in either a "plus" or a "minus" direction (that is, that the direction of the variable's influence was clear); that the variable measured part of a spectrum, or series of variables which, as a series, was worth retaining; or that the variable's scores were pooled, along with those of several other variables, to make up the scores of another variable which was a better discriminator.

The Important Groups of Variables. Table 6 summarizes the status of the six groups of variables at the beginning and end of the analysis. These groups, described earlier, were generalized categories based on a large number of individual variables. Most of the variables were represented in one or another of the groups. Table 6 shows the most important groups to be those of emotion, the personalities of the major characters, the plot themes, and simplicity. It must be recognized,

however, that one of these groups, that related to emotion, stood out above all the others in importance.

FINDINGS: THE FORMULAS

One of the purposes of the study was to find a combination of variables which would allow prediction of sale more accurately than could be done with one variable alone. Consequently, the sixteen variables in Table 5 were combined in various ways to ascertain which combination would correctly classify the highest percentage of the forty-four novels as best or poor sellers. The fewer variables used, the easier it would be to apply the formula. No attempt was made to predict the actual sale or even the *PW* score of a novel; a correct prediction meant only that a book was classified correctly as a best or poor seller.

The statistical technique used in developing the formulas was the discriminant functions technique. This was developed only a few years ago by R. A. Fisher to aid in separating the members of two groups on the basis of certain of their characteristics.²² It has previously been employed in botanical, anthropological, personnel, and financial work; applications have ranged from separating two species on the basis of petal length²³ or bone dimension,²⁴ to distinguishing between good and poor salesmen,²⁵ or between good and poor loan risks in the consumer installment finance business.²⁶

The Two-Variable Formulas. All of the sixteen variables were used at least once in the two-variable combinations. Readability (variable 1) was recognized as one of the best discriminators, so it was tried once with all other variables. This provided a comparison of the discrimination obtained with each variable; it pointed to certain variables and combinations as worth trying in additional formulas, and to others as of no value because of high intercorrelations.

Formulas with Three or More Variables. After the two-variable combinations had been worked out, three-variable combinations were tried. The variables showing up well in the two-variable formulas were tried in three-variable combinations in an attempt to raise the

²² R. A. Fisher, "The Use of Multiple Measurements in Taxonomic Problems," *Annals of Eugenics*, Vol. 7 (September, 1936), pp. 179-88; "The Statistical Utilization of Multiple Measurements," *Annals of Eugenics*, Vol. 8 (August, 1938), pp. 376-86.

²³ H. F. Smith, "A Discriminant Function for Plant Selection," *Annals of Eugenics*, Vol. 7 (November, 1936), pp. 240-50.

²⁴ M. M. Barnard, "The Secular Variations of Skull Characters in Four Series of Egyptian Skulls," *Annals of Eugenics*, Vol. 6 (December, 1935), pp. 352-71.

²⁵ N. Wallace and R. M. W. Travers, "A Psychometric Sociological Study of a Group of Speciality Salesmen," *Annals of Eugenics*, Vol. 8 (May, 1938), pp. 266-302.

²⁶ D. Durand, *Risk Elements in Consumer Installment Financing: Technical Edition* (New York: National Bureau of Economic Research, 1941).

level of successful prediction. Following the calculation of the three-variable formulas, four- and five-variable combinations were tried. They were developed on the same basis as the three-variable formulas; that is, from the best combinations discovered previously. The improvement of the four- and five-variable combinations did not seem sufficient to warrant trying any combination with more than five variables.

Of all the combinations worked out, Table 7 shows the best one (of those which did not require use of variables 15 and 16) to be the

TABLE 7
BEST COMBINATIONS

<i>Variables*</i>	<i>Percentage of Correct Prediction†</i>
1-4-15	82%
4-7-15	82
1-4-7-9-15	82
1-4-7-15	80
1-3-4-7-15	80
4-15	77
1-4-7-9	77
4-7-9-16	77
4-9-15	75
1-3-7-15	75

* Code

1. Readability.

3. Recency of the events in the novel.

4. Sentimental theme.

7. Central male character's affectionate attitude toward other characters.

9. Central male character's total of level 1 emotion.

15. Prior publication date.

16. Wordage of the novel.

† All discriminations in this table were significant at or below the 5 per cent confidence level.

1-4-7-9 formula (readability, central male character's total of level 1 emotion, an affectionate attitude by the central male character toward other major characters, and a sentimental theme). It will be noted that only one of these variables (central male character's total of level 1 emotion) had enough plus or minus scores to meet the 50 per cent criterion. The other three variables, which proved to be among the four best when used in combination, would have been eliminated if the 50 per cent criterion had been applied indiscriminately in the latter part of the analysis.

Testing the Significance of the Discrimination. For each formula it was necessary to test the hypothesis that its percentage of correct prediction was actually due to chance factors rather than a real and reliable ability to discriminate between best and poor sellers. For

the discrimination to be accepted as a real, rather than a chance discrimination, it had to meet the usual 95 per cent standard for level of significance, that is, there had to be at least 95 chances out of 100 that the discrimination was real and reliable. This test eliminated approximately a third of the combinations.

A second test of significance was designed to test the value of new variables which might be added to a formula. This test was intended to show whether or not the addition of a variable to a particular formula significantly increased the percentage of correct prediction. Several formulas were eliminated by this test.

Testing the Reliability of the Best Formula. In any statistical study, it is necessary to test the reliability of the results. The ideal method for this in the present study would have been to apply the formula to additional pairs of novels matched in the same manner as the original group, but the lack of good pairs which had not been used in the original group made it impossible to carry out this method. Themes, publicity factors, and sales totals would not have been matched as closely in the additional pairs as they were in the original pairs, and, consequently, the two groups—the original and the test groups—would not have been quite comparable. Therefore, another method of testing reliability, the split-half method, was employed.

In the split-half method, results from half of the original data are compared with results from the remaining half of the data; this measures the data's internal consistency and is a partial measure of reliability. In the present study, the reliability of the best content formula (the 1-4-7-9 formula) was tested by comparing the eleven "better matched" with the eleven "poorer matched" pairs of novels. The better matched pairs were those in which themes, *PW* scores, and publicity factors were closely matched, while the poorer matched pairs were those not matched quite so closely on these factors. A new formula was worked out with the data for the four variables from the eleven poorer matched pairs; this formula was called the "1-4-7-9 poor" formula. This new formula was then applied to all the novels. It classified correctly 77 per cent of the poorer matched pairs and 73 per cent of the better matched pairs. Since these scores were so close, we may safely assume that there was a high degree of internal consistency and reliability in the 1-4-7-9 formula.

CONCLUSIONS

In the Berreman study and the research here reported, two of the important areas of consideration—the best seller's publicity and its

content—have, for the first time, been studied carefully and correlated positively with sale. But while these studies have brought a certain amount of understanding of the problem, this understanding has not yet advanced beyond its elementary stages. It is not yet possible to state categorically just what the relations are between the conclusions of the two studies themselves. To understand their relationships it will be necessary to weigh the relative value of the factors they found important, and to find out when and under what conditions each is operative. Nor is it yet possible to state the conclusions of the two studies more than hypothetically for the novels they did not include. More evidence is needed before it will be possible to predict the effect on any one particular novel of its publicity and content factors.

The causal factors behind the sale of best sellers are sufficiently complex and are imbedded so deeply in the psychological and sociological aspects of modern culture that their description cannot be attempted, although the correlative factors found important in the two studies obviously suggest avenues of approach to the problem. Other areas which must be studied before all the correlative factors can be assembled are those relating to the author, the publisher, the "society," and the individual reader.

30. WHY DAILY DEPARTMENT STORE SALES FLUCTUATE*

Why do people shop on some days and not on others? To some extent because of the weather! That market behavior may be influenced entirely by physical phenomena is the principal point of this article.

Indeed it suggests the possibility, which has been substantiated in other areas, that weather conditions are responsible in part for one's psychological outlook, which in turn can lead to changes in buying habits. The article also advances some useful ways of incorporating weather variables in predictive relationships.

Many studies have been made of the effects of weather on various kinds of business activity in recent years. Studies have been made of weather's effect on gas consumption, electrical consumption, agricul-

* Adapted from an article by A. T. Steele, Mead Johnson and Co., "Weather's Effect on the Sales of a Department Store," *Journal of Marketing*, Vol. 15 (April, 1951), pp. 436-43.

tural production, electrical power failures, and so on. But little has been done on the subject of weather's effects on the retail sale of merchandise.

Weather affects gas consumption mainly through the effect of temperature on people. In cities in which gas is used for household heating, the relation between temperature and gas consumption is rather obvious. The relation between gas consumption and humidity and wind velocity is not so apparent, but there probably is one. With respect to electrical consumption, cloudiness as well as other weather elements affect consumption. When the sky is overcast and the clouds are thick, people will turn on their lights. Agricultural production is affected by many weather elements, but the time and amount of precipitation are probably the most important to most crops.

In the cases of gas consumption, electrical consumption, and agricultural production the relation with some one or more weather elements is rather evident. How weather affects the total sales of a retail department store is not so clear. It is conceivable, however, that weather might affect the sales of a retail store in the following ways:

1. The weather could be of such a nature that it is for one reason or another uncomfortable to go to the store.
2. The weather could produce situations that would physically prevent people from going to the store, as in the case of snow drifts over roads and streets.
3. Weather may have psychological effects on people that may change their shopping habits.
4. Some kinds of merchandise may be more desirable during a period in which certain types of weather prevail.

In this study,¹ no attempt was made to ascertain the extent to which one or the other of the above four factors affected sales. Weather was correlated with sales (on a daily basis) and the extent of correlation was calculated.

METHODOLOGY

Since sales and the weather can be expressed quantitatively, it was decided that a multiple regression equation relating sales and the weather should be developed. The dependent variable is daily store sales of three departments of Younker Bros., Inc., Des Moines, Iowa, and the independent variables are whatever weather variables are needed adequately to express the weather situation.

Rectification of Sales Data. Department store sales during the seven weeks before Easter (hereafter called the Easter Season) vary as

¹ A. T. Steele, *Weather as a Determinant of Marketing Strategy: A Department Store Case Study* (Ph.D. thesis, State University of Iowa, 1950).

between weeks and as between days because of shopping habits. Therefore, it was first necessary to correct the sales data for seasonal variation. The correction for daily and weekly variation was done by the percentage of moving totals method described in Croxton and Cowden, *Practical Business Statistics*.² Correction for trend was accomplished by dividing the resulting figures (that is, after correction for daily and weekly variation) by figures proportional to the average daily sales for the calendar year. The resulting figure was called "rectified daily store sales" or, more briefly, "rectified sales."³

Rectification of Weather Data. In the study, the following weather elements were used to describe the weather situation: (1) the amount of precipitation between 6:00 A.M. and 6:00 P.M.; (2) the depth of snow cover at 6:30 P.M.; (3) temperature at 12:30 P.M.; (4) wind velocity between 12:00 NOON and 1:00 P.M.; and (5) the amount of sunshine during daylight hours (as a percentage). Since the relationship between three of the individual weather variables and rectified sales is not linear, it was necessary to rectify⁴ them.

The snow-cover and sunshine data were not rectified; the other three were. Modifying (by statistical methods) the formulas defining the cooling power of the air developed by many students during the past thirty-seven years,⁵ a formula for rectifying the wind and temperature figures was developed. That formula was $C = V^{1/4} (80^\circ\text{F.} - T)$; in which C is an index of the cooling power of the air, V is the wind velocity between 12:00 NOON and 1:00 P.M., 80°F. is eighty degrees Fahrenheit, and T is the temperature at 12:30 P.M. in degrees Fahrenheit.

On the basis of studies by Professor Eugene Van Cleef⁶ and by Mr. Herbert Johnson,⁷ Sales Manager, Cushmans Sons, Inc., New York,

² F. E. Croxton and D. J. Cowden, *Practical Business Statistics* (2nd ed.; New York: Prentice-Hall, Inc., 1948), pp. 292 ff.

³ The term "rectification" is here used to denote the total process of correcting data for seasonal and trend variation. It follows the usage of J. Dean, "Department-Store Cost Functions," in O. Lange, F. McIntyre, and T. O. Yntema (eds.), *Studies in Mathematical Economics and Econometrics* (Chicago: University of Chicago Press, 1942), pp. 228 ff.

⁴ "... many of the functions that fit empirical data can be expressed, by a change of variable, in linear form. From the given table, a new table may be computed for the new variables, and the data in the new table are then said to be *rectified*." F. A. Ficken, "Meteorological Mathematics and Calculations," in F. A. Berry, Jr., E. Bollay, and N. R. Beers (eds.), *Handbook of Meteorology* (New York: McGraw-Hill Book Co., 1945), p. 232.

⁵ R. G. Stone, *On the Practical Evaluation and Interpretation of the Cooling Power in Bioclimatology*, Blue Hill Observatory, Harvard University, Reprint No. 10, 1943.

⁶ E. Van Cleef, "The Influence of Weather on Street Car Traffic in Duluth, Minnesota," *Geographical Review*, Vol. 3 (February, 1917), pp. 126-34.

⁷ Personal letter.

the precipitation figures were rectified by taking the logarithm of each figure.

Because of the changing date of Easter, two computations were made: (1) one based on the years in which the date of Easter was early, and (2) the other based on the years in which the date of Easter was late. Results were different mainly because early Easter seasons are typically colder and stormier than late Easter seasons.

Early Easter Seasons. The estimating equation (arrived at by the least squares method) relating the four weather factors to rectified sales is as follows:

$$S' = 104.73 - 0.0228C - 0.0219P - 0.2635N + 0.0675U$$

in which S' is rectified store sales as predicted by this estimating equation, C is the cooling power factor, P is the precipitation factor, N is the amount of snow on the ground at 6:30 P.M., and U is the per cent of sunshine.⁸

A calculation of the beta coefficients indicated that the amount of snow cover (N) was definitely the outstanding weather cause of sales variation. The amount of snow cover was the chief cause even though the snow fell several days or even a week before.

Late Easter Seasons. During three years that the date of Easter was late, the estimating equation relating the four weather elements to rectified sales is as follows:

$$S' = 109.51 - 0.0361C - 0.0429P - 0.2596N - 0.0189U$$

in which the symbols have the same meaning they were given above.⁹

The beta coefficients indicated that the precipitation factor had more effect on sales variation than the other three weather elements. Snow cover and cooling power had almost as much influence, however. The influence of sunshine was, as in the case of early Easter seasons, minor (though statistically significant at the 5 per cent probability level).

The Accuracy of the Estimated Equations. The accuracy of the results can be presented in several different ways. Figure 1 shows, in the form of a scatter diagram, the relation between the daily store sales

⁸Based on the Easter seasons of 1945, 1947, and 1948. There were 126 sets of data. Variance analysis indicated that all the independent variables made a significant contribution.

⁹Based on the Easter seasons of 1943, 1944, and 1946. There were 126 sets of data. Variance analysis indicated that all of the independent variables made a significant contribution at the 5 per cent level, though the contribution of the sunshine factor was not significant at the 1 per cent level.

as predicted by the appropriate estimating equation and the actual daily store sales. The figure shows a very high correlation between the actual sales and the sales as estimated by the regression equation. The Pearsonian coefficient of correlation between daily store sales as predicted by the estimating equation (for years with early Easter dates) and the actual sales figures was 0.94. Thus 88 per cent of the variance (0.94^2) of the sales of the store was accounted for by the weather variables considered and the rectifying process. This 88 per cent figure

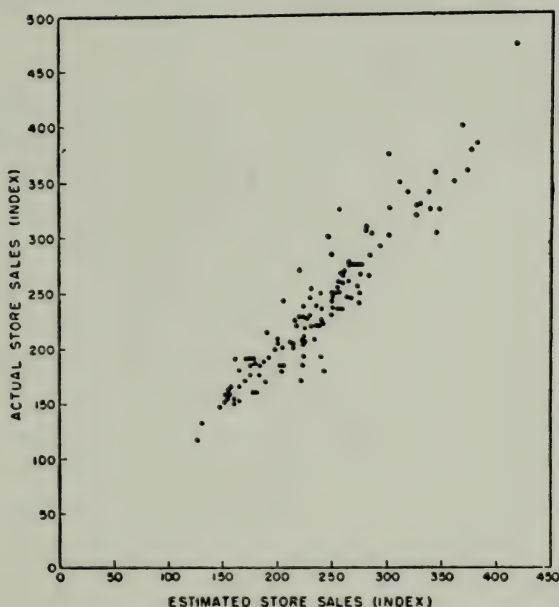


Fig. 1. Scatter diagram of the relationship between estimated store sales and actual store sales, Easter seasons, 1945, 1947, and 1948.

can be subdivided as follows: 42 per cent of the variance in the store's sales was due to the weather factors considered and the other 46 per cent accounted for by the rectification process used. An analysis of variance indicated that the results were statistically significant. The accuracy of the estimating equation for late Easter seasons was only slightly lower.

PRACTICAL APPLICATION OF THE STATISTICAL RESULTS

Three applications of the techniques here discussed are rather evident. The first and most apparent is the use of the estimating equations to forecast sales ahead of time by substituting the forecasted weather into the equations. Then the number of salespeople could be adjusted (the day before) to the predicted needs for salespeople. According to

Robinson and Brisco, a sales-force budget is expected to meet the following general requirements:

1. There should be a sufficient number of people to give adequate protection to the store's merchandise and fixtures.
2. There should be a sufficient number of employees to assure adequate customer service.
3. The number of people should be so planned that the group can accomplish the task set, and each individual can earn a reasonable wage.
4. The total wages paid to the salespeople should bear a percentage relationship to net sales commensurate with profitable past experience of the store . . .¹⁰

In general, these requirements when carefully applied should result in an almost constant ratio between salespeople's salaries and total dollar sales of the store. In 1947 the ratio of salespeople's salaries to dollar sales of all reporting department stores with annual sales over \$20,000,000 was 6.1 per cent.¹¹ The volume of sales varies (on the average) from hour to hour, from day to day, and from week to week. Adjustments in number of salespeople to these average variations is now common practice. But the average variation figures used do not, of course, "take cognizance of unusual weather conditions which might affect the actual volume achieved on a particular day."¹² Hence, the need for utilizing the probable effect of weather on sales when determining the number of "extras" or per diem salespeople to call in.¹³

A second application of the techniques developed is of an evaluative rather than a directive nature. The estimating equations can be used for the purpose of calculating daily sales potentials. For this purpose, actual weather observations (instead of forecasts) can be used. The calculated potential sales figure can then be compared with the actual sales figure for the day. The quality of the merchandising and promotional effort can then be evaluated. Weekly and monthly evaluations can be based on the sum of daily actual and potential sales figures.

Mr. H. C. S. Thom (formerly Section Director of the U.S. Weather

¹⁰ O. P. Robinson and N. B. Brisco, *Store Organization and Operation* (New York: Prentice-Hall, Inc., 1949), p. 180.

¹¹ Controllers Congress, 1947 *Departmental Merchandising and Operating Results of Department Stores and Specialty Stores*, National Retail Dry Goods Association, 1948, p. 97.

¹² M. P. McNair, C. I. Gragg and S. F. Teele, *Problems in Retailing* (New York: McGraw-Hill Book Co., 1937), p. 452.

¹³ In a like manner the amount of advertising might be varied to offset the effects of bad weather or to "make hay while the sun shines" (whichever policy seems best). In the study the results of an attempt to relate advertising lineage to sales and the pulling power of advertising to the weather were inconclusive. Much has yet to be done along this line.

Bureau in Des Moines) once related the effects of temperature on the truck route sales of a soft drink distributor in the Los Angeles area. The distributor used the results to evaluate the sales of each route by comparing them with the amounts that should have been sold on that route according to the temperature which prevailed during the heat of the day on that route. In this manner the distributor evaluated the performance of his salesmen truck drives.

A third application of the techniques discussed in this article concerns the relation between the store and its external economic environment to a greater extent than its internal control problems (although the two problems are interrelated, of course). The Federal Reserve Board's Index of Department Store Sales could be corrected for the effect of weather by applying the techniques here discussed to each store whose sales are included in that index. The estimating equations used above cannot be applied to the sales of any other store. But estimating equations can be developed for each store by the same methods.

If the Index of Department Store Sales which is already corrected for seasonal variation could also be corrected for the effect of weather, its usefulness should be greatly increased for the merchandise manager as well as for the student of business cycles. If the merchandise manager could know, for example, the department store sales fell, in the Federal Reserve District or city in which his store is located, by a certain number of percentage points because of unfavorable shopping weather, his buying decisions would be different than if the cause were primarily of an economic nature.

A NEW AREA FOR MARKET ANALYSIS

The techniques that have been used in this study can be applied to other seasons of the year and to other types of distributive organizations. Also, separate equations can be developed for each department in a department store.¹⁴ Practically every retail and wholesale establishment is affected by the weather. For example, large supermarkets in which the great bulk of sales are on Friday and Saturday are particularly interested in minimizing carry-over to Monday, and that carry-over is largely a function of weather conditions. Bakeries can greatly reduce stale bread losses by utilizing the method.¹⁵

¹⁴ In this study estimating equations were developed for the Candy Department and for the Moderate Priced Dress Department.

¹⁵ Cushmans Sons, Inc., of New York, saves an estimated \$250,000 per year by gearing production and distribution to the forecasted weather. If the weather is predicted accurately, Cushmans has very little day-old bread on hand. Daily total sales are

WEATHER FORECASTS

The greatest difficulty in the use of the method is in forecasting the weather one or more days ahead of time. (This is true only of the first application discussed above. The second and third applications do not depend on weather forecasts.) In general, any firm desiring to utilize the effect of weather on its daily sales to forecast the sales potential should contact a meteorological consultant.¹⁶ This meteorologist, in consultation with the firm's market analyst, should first make a complete statistical study of the exact effect of weather on that firm's sales in order to ascertain just what weather variables affect sales, how much, and whether or not the effect is different in different territories and at different times of the year.

The firm cannot rely on the U.S. Weather Bureau's general forecasts because they are entirely too general; Congress will not permit the Weather Bureau to make specific forecasts for private businesses (except those in transportation and agriculture).

Marketing research personnel will have some general knowledge of the effect of weather on the firm's operations. This general knowledge will be invaluable to the meteorological consultant. The meteorological consultant will have had at least some experience in general business and marketing problems. This knowledge too will make a major contribution. Together the meteorologist and the marketing analyst can do much to increase the efficiency of the firm.

31. WHY A MAIL-ORDER CATALOGUE IS SUCCESSFUL*

The success of Sears-Roebuck is a matter that has long intrigued merchandising people as well as business analysts generally. Many approaches have been taken over the years to identify the reasons for this success. None, however, is more novel than the research outlined in this article

affected by weather, but the greatest effect is on the location of bread purchases. People tend to buy more bread at downtown stores and less at suburban stores in bad weather. See "Weather Profit," *Business Week*, March 30, 1946, pp. 72-74.

¹⁶ For a list of meteorological consultants write the Executive Secretary, American Meteorological Society, 5 Joy Street, Boston 8, Mass., or to me.

* Adapted from an article by George Kingsley Zipf, formerly of Harvard University, "Quantitative Analysis of Sears, Roebuck and Company's Catalogue," *Journal of Marketing*, Vol. 15 (July, 1950), pp. 1-13. The author received constructive suggestions from Professor Reavis Cox of the Wharton School of Finance and Commerce.

to determine why people are attracted so consistently to the Sears catalogue.

The approach is an aggregate one from the seller's vantage rather than the buyer's. The reconciliation with buyer individual motivations appears to involve the implicit assumption that the structure of these personal motivations remains largely unchanged over time insofar as aggregate purchases (sales) are concerned. Though many may disagree with this point, and indeed with the meaningfulness of the findings presented in this article, the approach nevertheless is provocative and worth considering.

A stranger to our country who is interested in studying American habits of living, and has little time for his sojourn, would learn much from a careful study of one of the semiannual mail-order catalogues. The mail-order catalogue gives in compact form almost as complete an account of current American buying habits and consumer demands as one could ask.

Yet, to the serious student of marketing, the catalogue also represents a carefully considered selection of a wide diversity of articles offered in a variety of styles, in different sales units, and at different prices, which are displayed pictorially and verbally in a manner calculated to convince the reader both that the offered articles will effectively fulfill his needs, and that he needs the offered articles.

Insofar as the catalogue succeeds in thus convincing mail-order customers, the mail-order concern prospers. Insofar as the catalogue fails, however, the mail-order customers do not suspend their buying until the mail-order house has reconsidered its selections and issued a new catalogue. Instead, the public turns to competitors who have better guessed the public's needs and tastes.

For, after all, mail-order catalogues represent nothing more than the marketing guesses of the executives who are responsible for them, and who, if their guesses result in profits year after year, may be reckoned among the company's most valuable assets. Good guesses mean good profits; or, turned around, good profits represent the results of good guessing. Bad guessing, on the other hand, can be disastrous.

Although we should all like to be good guessers, we should scarcely agree on the explicit rules of good guessing. Nevertheless, we should all agree that certain procedures of guessing would be disastrously bad, and this agreement as to bad guessing may well provide the needed initial starting point for scientific research into good guessing—that is, into the principles of sound merchandising.

One procedure of admittedly bad guessing would be that of a sheer random selection, as would result, for example, if mail-order executives

solved all their merchandising problems by shaking dice, or by pulling numbers out of a hat, and the like. Or, the Sears executives could determine the number of different items to be offered in their next catalogue by dividing the treasury balance on a certain date by the average opening price of stocks on that date; the actual items to be offered could then be selected by lot from *Thomas' Register*.

That would in fact be a random selection which we should all agree would be a bad procedure because merchandising is not a random phenomenon. Yet if merchandising is not random, we agree by implication that merchandising is subject to the operation of underlying natural principles, even though no one may be able to state those principles precisely.

If there are principles of marketing, what are they? This is clearly a hard question. Nevertheless, in the light of the preceding argument, it is a meaningful question, and, therefore, a question that can be investigated with the empiric procedures of natural science, even though it may take years to do so.

There is no one single scientific procedure that will uncover the underlying principles of so highly complex a phenomenon as marketing. But there are certain scientific procedures that have been fruitful in the past in attacking similarly complex phenomena and which might be helpful here. For example, students of the principles of marketing might copy the scientific methods of the astrophysicists who, in studying the principles governing the behavior of the heavenly bodies—which they could not get into experimental laboratories—have no alternative but to observe carefully the actual behavior of the bodies in question, and then to seek to disclose any empiric regularities in the observed behavior.

After these two steps have been taken and empiric regularities have been disclosed, the next step is to inquire into the reason for the occurrence of the observed regularity instead of some other regularity, or instead of no regularity at all. That is, one asks: *why?*

This question, *why*, is of course a different kind of question from the *how* of behavior of the first two steps that preceded it. Indeed, we may never be able to agree *why* some things behave the way they do. But scientific methodology may well show *how* they behave; and that knowledge may have a tremendous practical value.

The above-described scientific method of seeking empiric regularities in historical phenomena is not new to serious students of marketing. It is essentially the method used by A. M. Wellington¹ in his study of

¹ A. M. Wellington, *The Economic Theory of the Location of Railways* (6th ed.; London: Chapman Hall; New York: John Wiley, 1906), pp. 713 ff.

transportation, by W. J. Reilly² in his pioneer studies of "retail gravitation," as subsequently elaborated and extended by P. D. Converse,³ by F. Strohkarck and K. Phelps,⁴ by J. Q. Stewart,⁵ and others, including myself who, in a recent book,⁶ presented some hitherto unobserved empiric regularities. In presenting the empiric regularities, a theory was advanced to suggest an answer to the *why* of the regularities reported.⁷ Mention of this theory is made here for two reasons. First, the theory led deductively to the disclosure of the empiric regularities about to be presented; hence the theory may have some practical value in helping to disclose still further empiric regularities. Second, the theory tenders a possible answer to the *why* of the empiric regularities I am about to present—an answer that limitations of space preclude repeating again in the present article.

EMPIRIC REGULARITIES IN THE SEARS CATALOGUE

The Price-Diversity Ratio. When one opens the Sears Catalogue, it is at once evident that prices vary quite widely for the different products offered, from hundreds of dollars for some items to a few cents for others. A question naturally arises as to whether there is any observable regularity in the distribution of the different items listed at the various prices falling within the price range of the entire catalogue.

For example, as the price increases, does the number of different items offered at that price decrease, increase, or behave in any other regular way? This problem is somewhat difficult because of the differences in price for different sizes of the same item. Thus, for example "household type" paint brushes "for every budget" (p. 830 in the Boston, Massachusetts edition, which was used throughout) come in seven different sizes, and at seven different prices, ranging from 15¢

² W. J. Reilly, "Methods for the Study of Retail Relationships," University of Texas Bulletin, No. 2944, 1929.

³ P. D. Converse, "A Study of Retail Trade Areas in East Central Illinois," *University of Illinois Bulletin*, Vol. 41, No. 7, Business Studies No. 2 (October, 1943); also his "Retail Trade Areas in Illinois," *ibid.*, Vol. 43, Business Studies No. 4, 1946.

⁴ F. Strohkarck and K. Phelps, "The Mechanics of Constructing a Market Area Map," *Journal of Marketing*, Vol. 12 (April, 1948), pp. 493-96.

⁵ J. Q. Stewart, "Empirical Mathematical Rules Concerning the Distribution and Equilibrium of Population," *Geographical Review*, Vol. 37 (July, 1947), pp. 461-85; also his "Demographic Gravitation: Evidence and Applications," *Sociometry*, Vol. 11 (1948), pp. 31-57.

⁶ G. K. Zipf, *Human Behavior and the Principle of Least Effort* (Cambridge, Mass.: Addison-Wesley Press, 1949).

⁷ *Ibid.*, Chap. 9, "The Economy of Geography," and Chap. 12, "Prestige Symbols and Cultural Vogues."

through \$2.30. Which price is selected? For convenience, we have selected the median price (55¢) for this particular paint brush, and have done likewise for all other items offered at different prices for different sizes. On the other hand, when an item was offered in different qualities and styles, the price of each different style and quality was treated separately as if it were a different item.

After these decisions had been made, the first step, which was undertaken with the generous help of my students, was to ascertain the number of different items in the catalogue of the same price in price intervals of 25¢ up through \$2.00; then in 50¢ price intervals up through \$5.00; then in \$1.00 price intervals up through \$20; then in price intervals of \$5.00 up through \$100; and then in price intervals of \$25 up to the top of the price range. For example, 775 different items were offered at 25¢ or less; 1,042 different items were offered in the 50¢ price interval from \$2.00 through \$2.50; and only one item was offered in the \$25 price interval from \$400 through \$425.

The second step was to find the number of different items at the class midpoints in units of 25¢ for those price intervals that were larger than 25¢. Thus, for example, the 50¢ price interval from \$2.00 through \$2.50 in which 1,042 articles were quoted was considered to consist of two 25¢ price intervals; hence, the 1,042 different articles were divided by 2, and the dividend, 521, was assumed to be the number of items at \$2.25, which was taken as the class midpoint between \$2.01 and \$2.50. As another example, the \$25 price interval between \$400 through \$425 in which only one item was quoted, was considered to consist of one hundred 25¢ price intervals, with the class midpoint at \$412.50 at which .01 item was assumed to be offered. Since there is no such thing as .01 of an item, we shall understand .01 as the average number of items found at the class midpoint.

The third step was to graph the data on double logarithmic coordinates, with price (in units of 25¢) measured vertically on the y -axis, and with the average number of items at each class midpoint measured horizontally on the x -axis. The data, thus graphed and presented in Figure 1, seem by and large to descend from left to right in a straight line.

The fourth step was to calculate by least squares the best straight line through these points. This calculated line is drawn through the points of Figure 1. The equation for this line is $\log y = -0.5517 \log x + 3.9098$, with the error, ± 0.2475 .

In view of the above equation and error, we may say that *the number of different items of like price tends to vary in inverse proportion to the*

square of the price. In short, we have here an "inverse square" relationship.⁸

Although the general rectilinearity of the distribution of Figure 1 is unmistakable, it is evident that the 3 or 4 points at the bottom and a few at the top fall below the calculated line. Since the Sears management alone decides the price range of its goods, we may ponder the reason for these slight deviations at the extremes of the distribution.

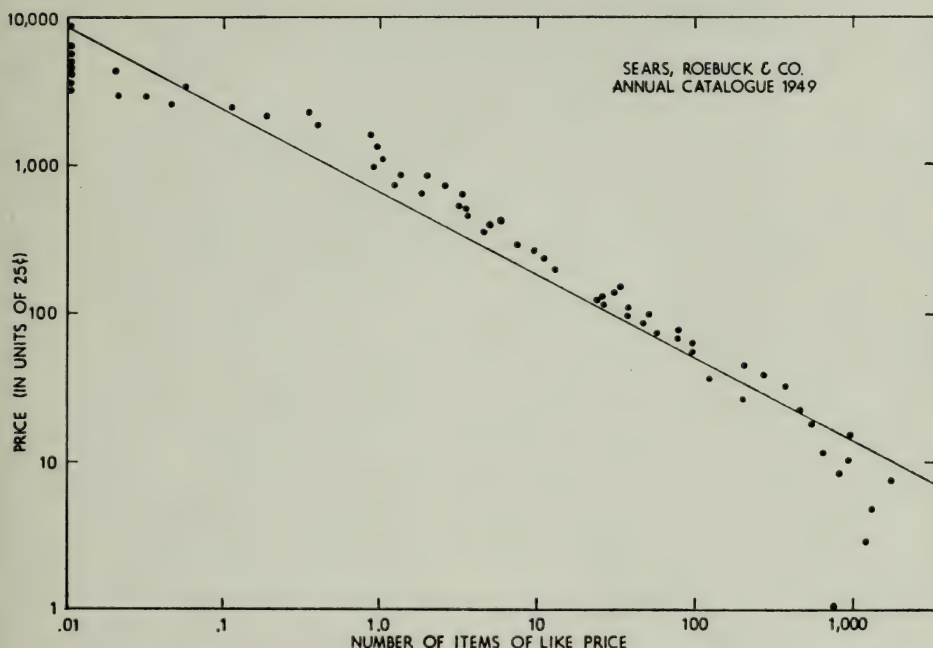


Fig. 1. The distribution of items of like price (unit: 25¢). (Sears, Roebuck Catalogue, Spring-Summer Edition, 1949.)

The fact that the lower points, representing items of about \$1.00 and less, fall unmistakably below the line may mean that persons who need the items in this lower price range tend to buy from local variety stores wherever possible instead of bothering to order them by mail, unless they happen to be ordering other items. Moreover, there is presumably little incentive for the mail-order house to expand its line of low-cost goods, except as a convenience for customers in outlying districts, because of the small profit in these goods after filling the individual orders for them.

⁸ This particular distribution of "the inverse square" which is corollary to the harmonic distribution was selected for graphical presentation, instead of the harmonic distribution, because of the greater ease of calculating the slope. For mathematical relationship between the two distributions, see G. K. Zipf, *op. cit.*, p. 547, Ref. 10 to Chap. 2.

The deviation of the very high-priced goods below the line at the top left may mean that a mail-order house operates at a disadvantage in the high-priced range because customers of high-priced goods may desire to inspect the actual goods before making so large an outlay for the purchase. This inspection is impossible with a mail-order house unless it has a retail store fairly close at hand that stocks the item in question. Hence, the management might understandably be cautious about increasing its line of high-priced goods; and this caution may well explain the deviation below the line at the upper left. The fact that this deviation does not appear until the price is well up into the hundreds

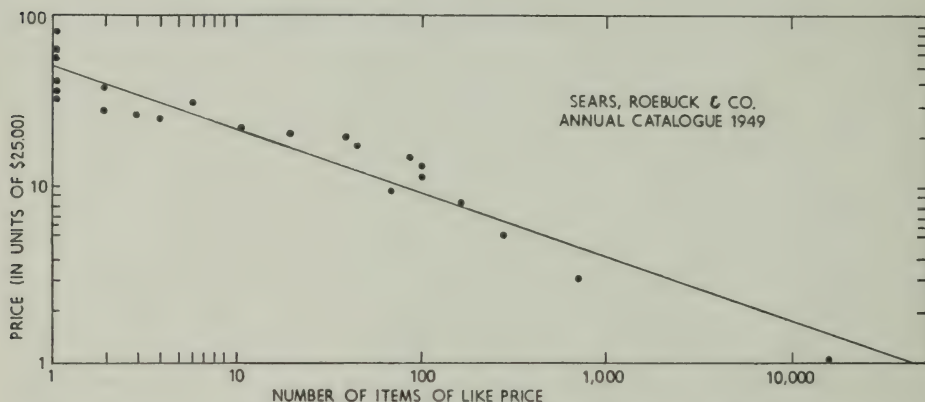


Fig. 2. Distribution of items of like price (unit: \$25). (Sears, Roebuck Catalogue, Spring-Summer Edition, 1949.)

of dollars (for most persons, quite large sums) would seem to indicate the confidence the firm enjoys in the public mind, in the sense that the public tends to accept the catalogue's word about its goods, unless the sums involved are very large.

The variation of Figure 1 would be decreased if a larger price interval were selected. Although, at the present preliminary stage of this research project, the essentially rectilinear nature of the distribution is of primary interest, nevertheless we present for their didactic value the same data in Figure 2, with the price interval increased 100 fold to \$25. In spite of this very large price interval, the general rectilinearity of the distribution remains, even though the slope, as calculated by least squares, decreases to -0.3678 ± 0.1141 (the line having the equation, $\log y = -0.3678 \log x + 1.7064$).

In view of the distributions of Figures 1 and 2, which are by no means an *a priori* necessity, and which are scarcely to be expected if the Sears management made its decisions by random selections, we can

only conclude that some sort of natural law is operating in the price-diversity ratio.

At this point several questions are in order.

First, does this observed price-diversity ratio in the Sears Catalogue merely reflect the price-diversity ratio of goods throughout the country? Theoretically, yes.⁹ Actually we do not know (that is the point of the research project).

Second, will this price diversity ratio be found only in stores that offer a wide diversity of goods? Theoretically, it will diminish in stores that are more specialized.

Third, since the data are unweighted for sales volume, how would a weighting by sales volume influence the distribution? Theoretically, sales volume is inversely proportional to the square of the price;¹⁰ for this proposition there is some empiric support from confidential data. The distributions of Figures 1 and 2 refer to prices of goods quoted and not to prices of goods sold.

Fourth, what has all this to do with the sizes of sales units in terms of which prices are quoted? Actually, a great deal, since an alteration of sales unit with corresponding alteration of price could drastically alter the resulting price-diversity frequency distribution. Thus, tacks sold only in sales units of a ton, or paint by the drum, or textiles by the square rod would fall differently on the frequency curve. The fact that the distribution of Figure 1 is in terms of the familiar conventional retail sales units of the most heterogeneous sorts (volume, weight, length, area, "assortments," numbers, and so forth) means that these sales units, though for the most part conventional, represent the results of the operation of some underlying natural principle that presumably governs what particular sales units are convenient for customer and seller in the retail trade, since a virtually endless number of other sales units are *a priori* possible.

The foregoing questions are asked primarily to suggest that many other marketing factors may be closely related to this one striking empiric regularity just reported, that seems to lie at the very core of marketing.

The Style-Product Ratio. Some products listed in the Sears Catalogue are offered in many different styles, including colors, while others are offered in a few or only one. The question arises as to the possibility of some underlying regularity in the phenomenon of styles

⁹ *Ibid.*, Chap. 9, pp. 347-74.

¹⁰ *Ibid.*, pp. 371 ff.

—a question that potentially involves the philosophically moot problem of deciding whether we are dealing merely with a different style of the same product or with a totally different product.

In the Sears Catalogue, it is by and large not difficult to discriminate between styles and products because of the very way that the different styles and products are presented. Sometimes the different styles are enumerated (for example, p. 3, "5 styles at \$2.88," "4 styles at \$3.88"). Sometimes the difference in style is indicated by different letters; thus, for example, on page 391 there are seven different styles of men's shoes all offered at the same price yet differently lettered. The different patterns (that is, styles) of Axminster broadloom (pp. 593-95) are lettered from A through S. The same applies to fluorescent lights, bed spreads, pipes, and children's hats. In those few cases where the difference in style is not actually indicated, a group of persons, acting as a jury of experts, may on the whole tend to agree as to what are different products, and what are merely different styles of the same product.

For that matter the present statistical investigation of the style-product ratio in the catalogue was undertaken in such a way as, in a sense, to appeal to what might be called a jury of experts (the students working on the study). The catalogue was divided into four dozen samples of equal length with each student receiving a sample.¹¹ Each student examined his sample page by page and determined for each page the number of different styles per product (hence "styles per product per page"). Thus, on page 3, there is one product with four styles and one with five (by and large the different styles for a given product appear on the same page). After the students had completed their tabulations of the number of different products with the same number of different styles—that is, the number of products that had one style, the number that had two styles, the number that had three styles, and so on up—the statistical information of the four dozen samples was combined into a single tabulation for the entire catalogue. An inspection of the underlying individual samples disclosed essentially the same type of distribution in the samples as in the whole, except, of course, for the degree of variation.

The combined data for the entire catalogue are presented on doubly logarithmic co-ordinates in Figure 3, with the number of different styles per product plotted on the y -axis, and with the number of different products with the same y -number of styles on the x -axis. The

¹¹ The same procedure was used for all the other analyses reported in this paper, except for the index to the catalogue.

general rectilinearity of the distribution is unmistakable. The line drawn, which was fitted by least squares through all the points, has a -0.5593 slope (or, in equation form, $\log y = -0.5593 \log x + 1.9813$; ± 0.1859).

In view of the above equation and error, we may say that the *x-number of different products for which any y-number of different styles is listed tends to vary in inverse proportion to the square of y—again the “inverse square.”*¹²

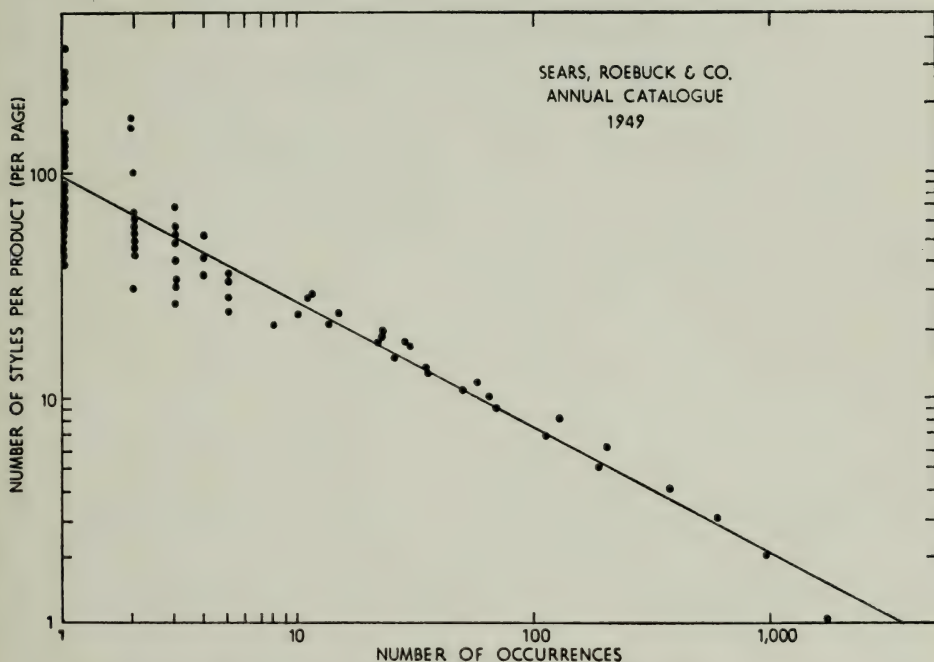


Fig. 3. Distribution of styles per product (per page). (Sears, Roebuck Catalogue, Spring-Summer Edition, 1949.)

The Display of Products. Even after the mail-order house has decided upon the kinds of merchandise it will offer, and its styles, unit prices and the like, it still has the task of presenting the goods effectively in the catalogue. Since catalogues cost money, we can understand the tremendous urge to condense the description of each item as far as possible. On the other hand, inasmuch as larger page areas are more likely to catch the eye than smaller ones, we can understand the urge to devote larger areas to the description of those items to which attention is to be particularly attracted. Hence we may reasonably expect to find a variation in area in the descriptions, or displays, of goods.

¹² See Zipf, *op. cit.*, Chap. 12, "Prestige Symbols and Cultural Vogues" for discussion of dynamics of style changes.

One method of assessing the comparative area devoted to the different products in the catalogue is to count the number of different products (ignoring as irrelevant all differences in product sizes) that are listed on each page, as determined by the actual number of different prices listed on the page. Under this procedure, when different styles are listed and priced separately, they are considered different products for the purposes of this part of the study, which is concerned only with the relative amounts of display area per product. Obviously, as the number of products per page increases, the average display area per product on that page decreases correspondingly.

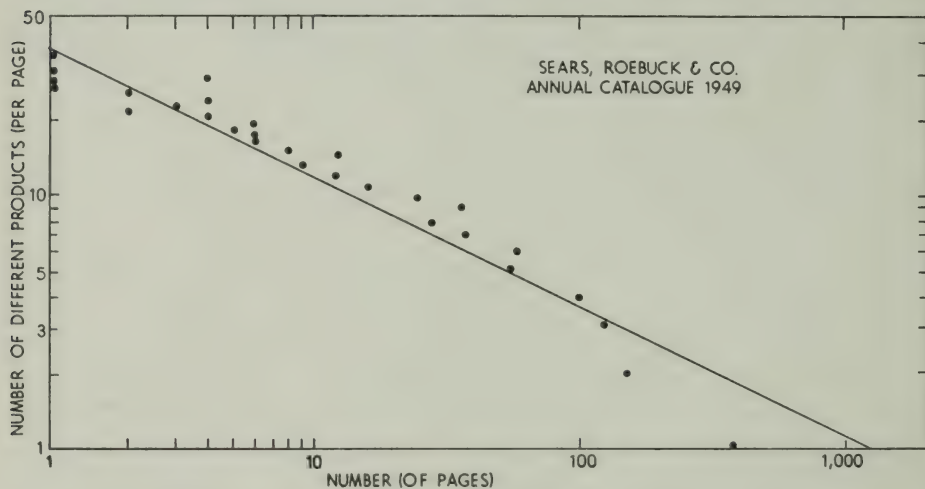


Fig. 4. Distribution of products (per page). (Sears, Roebuck Catalogue, Spring-Summer Edition, 1949.)

Then, by classifying the 1,100 odd pages of the catalogue according to the number of products listed per page, we shall be able to note any regularity in the use of display area.

In Figure 4 are presented on doubly logarithmic co-ordinates the x -number of different pages with the same y -number of different products. Except for the bottom two points (for pages with only one or two products), the distribution may fairly be viewed as rectilinear. The line drawn which was calculated by least squares through all the points has a -0.5084 slope (or, in equation form, $\log y = -0.5084 \log x + 1.5840; \pm 0.0787$).

Hence, *the x -number of different pages with the same y -number of different products tends to vary in inverse proportion to the square of y ; again the "inverse square."*¹³

¹³ See Zipf, *op. cit.*, pp. 170-78, and 564, Ref. 38 and Chap. 5 for the same type of distribution in other kinds of published materials, e.g., news-items, advertisements. See

Another method of gaining insight into the question of the relative display of products is to count the number of different pictures per product on each page (that is, "pictures per product per page") in the sense that every picture is to be counted as one unless it refers to precisely the same product, or unless it illustrates some special part of the same product, or the use of the same product. Thus, in the upper right hand corner of page 1083 is the picture of a "new farm-master estate sprayer" with a second picture that shows a man using the sprayer on a tree: together, one picture. The two baby chicks pictured together on

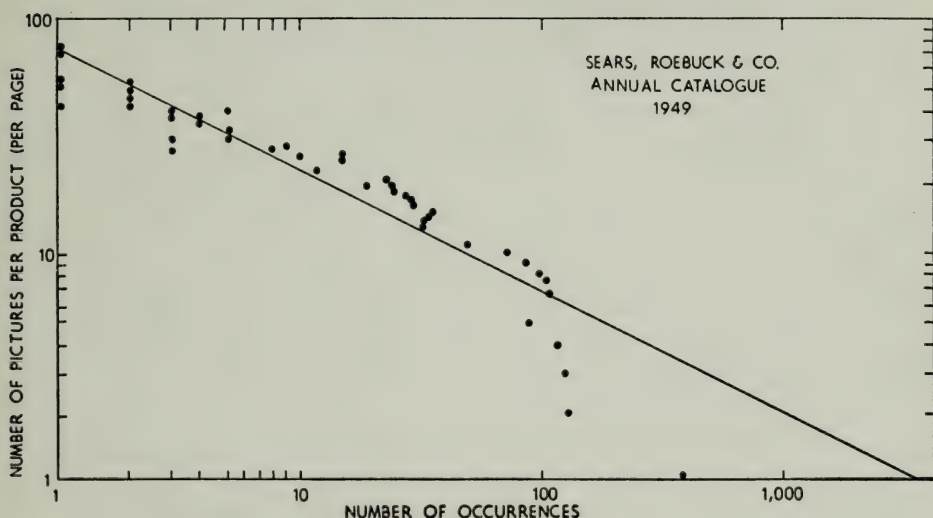


Fig. 5. Distribution of pictures (per product per page). (Sears, Roebuck Catalogue, Spring-Summer Edition, 1949.)

page 1112 represent one picture. On page 669 are thirteen different pictures referring to the whole and parts of the single "Kenmore Master DeLuxe Range"; one picture. Although the problem of classifying different pictures as above described may seem to be difficult, it admits in fact of little indecision, as will be noted from an actual inspection of the catalogue. Cartoons or other pictures that are occasionally added for purposes of atmosphere but which do not depict a specific product were ignored.

In Figure 5 are presented on doubly logarithmic co-ordinates the x -number of different pages that contained the same y -number of pictures per product for the entire catalogue. Except for the bottom 4 or 5 points, which indicate an avoidance of pages with very few pictures,

also W. B. Bryant, S. L. Washburn, and D. G. Outerbridge, "Some Psychological Determinants of the Structure of Advertising in a Classified Telephone Directory," *American Journal of Psychology*, Vol. 61 (October, 1948), pp. 540-44.

the distribution is rectilinear. The line drawn, as calculated by least squares for all the points, has a -0.5310 slope (or, in equation form, $\log y = -0.5310 \log x + 1.8685; \pm 0.1405$).

In view of this equation we may say that *the x-number of different pages with the same y-number of different pictures tends to vary in inverse proportion to the square of y; again the "inverse square."*

Hence, the distribution of pictures of Figure 5 tells us substantially the same as that of products of Figure 4. *A fortiori* there is a correlation between the number of pictures and number of products per page, as is clearly evident from an inspection of the catalogue, although an actual correlation study has not been made. In the overwhelming number of cases, products are pictured.

The Index. It is quite likely that the above unit of one single page is not suitable for all problems relating to the structure of the Sears Catalogue. Indeed, if instead of having selected the above problems of display that referred to individual products and styles—such as individual tree-sprayers, or particular styles of men's shirts—for which the unit of a page suffices, we had selected a more generic classification of merchandise, such as farm equipment or men's clothes, we should have found that in many cases the items of the more generic classes would have extended over many pages. In that case we should have measured the classes of goods in terms of the number of different pages over which they extended. Thus, according to the catalogue's index, farm equipment is displayed on a total of 65 different pages, whereas men's clothes are displayed on 50 different pages.

In a very practical sense the number of different items listed in the catalogue's 16-page index, when classified according to the number of different pages to which they refer, will give information about the relative amount of space devoted to different classes of items. This information on the index is presented in Figure 6.

In Figure 6 is plotted on doubly logarithmic co-ordinates the x -number of different items that are displayed over the same y -number of pages. The distribution is unmistakably rectilinear. The line drawn, which was calculated by least squares for all the points, has a -0.4238 slope (or, in equation form: $\log y = -0.4238 \log x + 1.6055; \pm 0.0640$). This slope of -0.4238 ± 0.0640 is a significant deviation below the previous -0.5000 slopes which have been observed for other indices.

This deviation below a -0.5000 slope is to be expected inasmuch as the same items are often referred to more than once in the index under different classifications (for example, "see *clothes* and *coats*"). In

other words, the more generic and the more specific classes have been combined into one index. Upon reflection, we can see that this combination might well have the effect of extending the entire distribution to the right, thereby decreasing its slope.¹⁴

Yet our concern for the niceties of slope should not disguise the remarkable rectilinearity of the distribution of Figure 6, particularly for the lower dozen or so points. This rectilinearity clearly indicates the

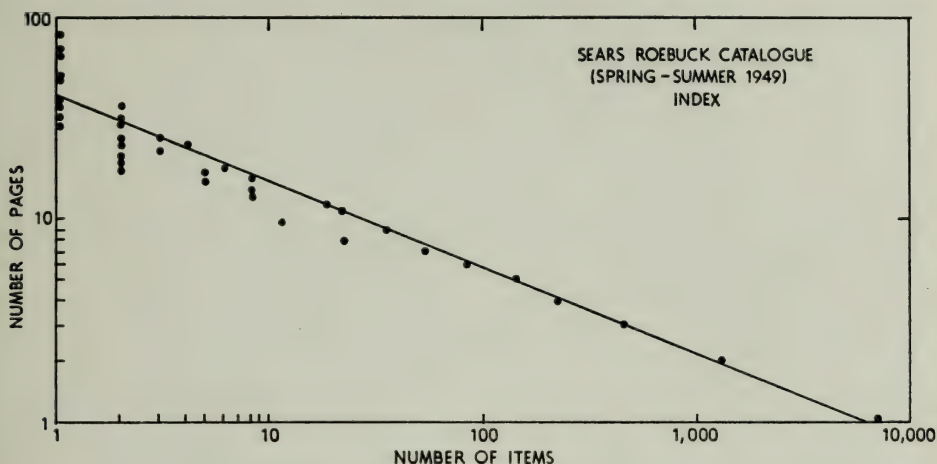


Fig. 6. Distribution of items in index. (Sears, Roebuck Catalogue, Spring-Summer Edition, 1949.)

presence of an underlying regularity both in the organization of the index in particular and of the catalogue as a whole.

DISCUSSION

The above distributions, which are scarcely due to the operation of sheer random chance—or “random selections”—represent empiric laws of a high degree of regularity in merchandising, pricing, and advertising of a very wide range of products. Moreover, they suggest that these different functions are organized together as a unit. The distributions do not represent principles that the executives follow consciously with calculating machines and slide-rules. Rather, they represent what may eventuate when, as in the present case, management is very efficient.

¹⁴ My former student, D. D. Bourland, Jr. has already (*and previously*) observed this slope in the indices of books (privately communicated to me, and not yet reported); also observed in the index to the *Congressional Record* (74th Congress), reported Zipf, *op. cit.*, p. 539, n. In the *Index to Thomas' Register of American Manufacturers* (37th ed.; New York: Thomas Publishing Co., 1946), Vol. 4, the x -number of generic headings that have the same y -number of specific subclasses is inversely proportional to y^2 (to be reported subsequently).

As empiric laws the distributions seem to have a two-fold value. First, they have a value for "pure" social science; indeed, in the present case they provide an empiric test of a general thesis already advanced.¹⁵ Second, they have a value for "applied" social science, in the sense that they may lead to the disclosure of fundamental principles that will be of practical value for the producers and distributors in our national economy. Further research on mail-order houses and related enterprises is in process.¹⁶ We have not presented the present study in the belief that it gives the entire picture.¹⁷

¹⁵ See Zipf, *op. cit.*, Chap. 9.

¹⁶ Considerable research has already been completed on the organization of *Thomas' Register*, *op. cit.*, and on the topic of brand names, both as listed *ibid.*, Vol. 3, and as studied in pantry polls and consumer analyses. Some of the observations already made, and which will be reported in a subsequent publication are: (1) the x -number of different manufacturers of the same y -number of products is inversely proportional to y^2 ; (2) the x -number of different concerns that have the same y -number of branch offices, warehouses, district sales offices, etc. is inversely proportional to y^2 ; (3) the x -number of different brand-names (including trade-names) that are owned by the same y -number of different firms is inversely proportional to y^2 ; (4) the x -number of different brand-names for the same y -number of different products is inversely proportional to y^2 ; (5) the same "inverse-square" relationship applies to the x -number of different brand-names that are used by the same y per cent of families; (6) the y -number of different brand-names in a city of P -size varies with the square root of P ; etc., etc., etc. Reported, G. K. Zipf, "A Note on Brand-Names and Related Economic Phenomena," *Econometrica*, Vol. 18 (July, 1950), pp. 260-63; also G. K. Zipf, "Brand-Names and Related Social Phenomena," *American Journal of Psychology*, Vol. 63 (July, 1950), pp. 342-66.

¹⁷ Thus, for example, there is a further point about a mail-order catalogue that seems to merit investigation: the ordering of items from beginning to end. A mail-order catalogue is essentially a department store in two dimensions. How far is its arrangement of items in the probable order of meeting of the reader's eye analogous to the arrangement of the items of a three dimensional department store, which also has the problem of display and of arranging items in reference to the probable order of meeting the customer's eye?

Bibliography

GENERAL REFERENCES IN MOTIVATION RESEARCH

ABT, L. E., and BELLAK, L. (eds.). *Projective Psychology*. New York: Alfred A. Knopf, 1950.

A series of articles by specialists in a variety of areas related to projective psychology. This provides an excellent summary of literature on the theory and practice of projective methods.

ADVERTISING RESEARCH FOUNDATION. *A Bibliography of Theory and Research Techniques in the Field of Human Motivation*. New York, 1956.

In most instances a brief annotation is given for each reference in the bibliography that includes the major disciplines, such as economics, sociology, anthropology, political science, and psychology, as well as marketing and advertising.

ALDERSON, W. *Marketing Behavior and Executive Action*. Homewood, Ill.: Richard D. Irwin, 1957.

An excellent treatment of the way in which group behavior (as described in such varied fields as sociology, psychology, anthropology, and political science) contributes to an understanding of marketing behavior.

———. "Survival and Adjustment in Organized Behavior Systems," pp. 65–88 in *Theory in Marketing*, R. COX and W. ALDERSON (eds.). Homewood, Ill.: Richard D. Irwin, 1950.

A discussion of behavior systems and their application to the study of market behavior.

ALLPORT, F. H. *Theories of Perception and the Concept of Structure*. New York: John Wiley and Sons, Inc., 1955.

Various approaches to the theory of perception and their application to the methodology of science are given. This will serve as a background reference for those interested in some of the technical literature related to motivation research.

ANDERSON, H. H., and ANDERSON, G. L. (eds.). *An Introduction to Projective Techniques*. New York: Prentice-Hall, Inc., 1951.

A thorough, almost encyclopedic, coverage of the field of projective techniques. Each author is a leading proponent of the particular technique.

BECK, S. J. *Rorschach's Test*, Vols. I and II. New York: Grune & Stratton, Inc., 1951.

A complete coverage of processes used in administering the Rorschach test is given as disproof of the contention of certain intuitionists concerning the essential incommunicability of their skills.

BELL, J. E. *Projective Techniques*. New York: Longmans, Green, & Co., Inc., 1948.

This book provides a comprehensive review of the literature on projective techniques to 1947, a manual for the use of projective techniques, a stimulus to research workers to become more familiar with these areas, and a discussion of related techniques. Extensive bibliographies for further reference are included.

BLAKE, R. R., and RAMSEY, G. V. *Perception—An Approach to Personality*. New York: Ronald Press Co., 1951.

An excellent symposium on the perceptual problems related to the multiplicity of behavior pathologies, unconscious processes, and psychotherapies which a number of authors contributed as papers in this book.

BONNER, H. *Social Psychology, an Interdisciplinary Approach*. New York: American Book Co., 1953.

An excellent book for marketing and motivation research workers because of the thorough way in which it covers such major topics as the social matrix of behavior, cultural values and personal-social adjustment, social change and collective behavior, and group tensions and culture.

FRANK, L. K. *Projective Methods*. Springfield, Ill.: Charles E. Thomas, 1948.

A general discourse on the nature of projective methods and means by which they can be used to advance the study of personality. It is accompanied by an extensive bibliography.

RORSCHACH, H. *Psychodiagnostics*. Berne, Switzerland: Hans Huber, 1942. New York: Grune & Stratton, Inc., 1942. (Translation and English edition by P. LEMKAU and B. KRONEBERG; editor, W. MORGENTHAUER.)

This book was prepared from a translation of Rorschach's original and basic work on the Rorschach test. It gives an excellent over-all view of his work for those who are interested in the technical aspects of his clinical techniques.

SMITH, G. H. *Motivation Research in Advertising and Marketing*. New York: McGraw-Hill Book Co., Inc., 1954.

A simple, well-written introduction to basic concepts in motivation research, with heavy emphasis on projective techniques and their application to commercial research.

WALES, H. G., and FERBER, R. *A Basic Bibliography on Marketing Research*. Chicago: American Marketing Association, 1956.

An interdisciplinary approach is taken in this bibliography, prepared for those interested in the literature related to marketing research and motivation research.

WHITE, R. W. *The Abnormal Personality*. New York: Ronald Press Co., 1956.

This is a descriptive book on the clinical level of the abnormal personality. Its value to marketing and motivation research workers is principally in the varied approaches that have been given to numerous behavior problems.

WULFECK, J. W., and BENNETT, E. M. *The Language of Dynamic Psychology, as Related to Motivation Research*. New York: McGraw-Hill Book Co., Inc., 1954.

This is a glossary of the terms used in or related to the field of motivation research. Each term is defined and then a short explanation is given.

PART I. GENERAL APPROACHES

1. A PSYCHOLOGIST'S VIEW BY CHARLES CANNELL

No references listed.

2. ADVERTISING STRATEGY AND THEORIES OF MOTIVATION BY WROE ALDERSON

ROGERS, C. R. *Client-Centered Therapy*. Boston: Houghton Mifflin Co., 1951.

A discussion of a current view of client-centered therapy and its application to a variety of situations. Potential applications for marketing research workers can be found in Part II.

3. TOWARD AN UNDERSTANDING OF HUMAN BEHAVIOR BY ERNEST DICHTER

COHEN, M. R., and NAGEL, E. *An Introduction to Logic and Scientific Method*. New York: Harcourt, Brace & Co., 1942.

A standard reference work on logic and scientific method.

MERTON, R. K. *Mass Persuasion*. New York: Harper & Brothers, 1946.

During World War II the bond drives presented an opportunity to try out techniques of mass persuasion. This book points out how a thematic analysis for the drives was used and the ways in which it was applied.

PUMPIAN-MINDLIN, E. (ed.). *Psychoanalysis as Science*. Stanford: Stanford University Press, 1952.

The problems and techniques of psychoanalytic validation and progress are discussed fully in this reference book.

4. HOW PSYCHIATRIC METHODS CAN BE APPLIED TO MARKET RESEARCH BY JAMES M. VICARY

No references listed.

5. A CASE STUDY: WHAT AUTOMOBILES MEAN TO AMERICANS BY PIERRE D. MARTINEAU

No references listed.

6. MOTIVATION RESEARCH—OPPORTUNITY OR DILEMMA? BY ALFRED POLITZ

ANDERSON, H. H., and ANDERSON, G. L. (eds.). *An Introduction to Projective Techniques*. New York: Prentice-Hall, Inc., 1951.

See earlier reference in this bibliography under General References in Motivation Research, page 403.

GLOVER, E. *The Technique of Psycho-Analysis*. New York: International Universities Press, Inc., 1955.

A new edition of a text that was originally written in 1928. It is intended for clinical psychologists and psychiatrists and will have limited value to the layman.

HOLTZMAN, W. H., and SELLS, S. B. "Prediction of Flying Success by Clinical Analysis of Test Protocols," *Journal of Abnormal and Social Psychology*, Vol. 49 (October, 1954), pp. 485-90.

These tests indicate the possibilities of developing a battery of personality and motivational measures that predict adjustment to military and combat flying.

ZUBIN, J. "Failures of the Rorschach Technique," *Journal of the Projective Techniques*, Vol. 18 (September, 1954), pp. 303-15.

A critical examination is made of the failures of the Rorschach technique with suggestions on ways to clarify many of the present-day contradictions and to obtain a better perspective on the evaluation of personality.

7. RATIONALITY AND IRRATIONALITY IN MOTIVATION RESEARCH BY L. EDWARD SCRIVEN

No references listed.

PART II. PROJECTIVE AND RELATED PSYCHOLOGICAL TECHNIQUES

8. PRINCIPAL PSYCHOLOGICAL TECHNIQUES IN MOTIVATION RESEARCH BY LYNDON O. BROWN

No references listed.

9. SOME HAZARDS OF MOTIVATION TECHNIQUES BY STEUART HENDERSON BRITT

No references listed.

10. AN APPLICATION OF PROJECTIVE TECHNIQUES BY MASON HAIRE
No references listed.

11. PROJECTIVE TECHNIQUES IN ATTITUDE MEASUREMENT BY IRVING R. WESCHLER AND RAYMOND E. BERNBERG

ALLPORT, G. W. *The Personal Document in Psychological Science*. New York: Social Science Research Council, Bulletin No. 49 (1942).

The results that can be achieved by the use of personal documents (life histories) of subjects is excellently reported in this bulletin. It shows how the self-revealing record (life history) intentionally or unintentionally yields information regarding the structure, dynamics, and functioning of the author's mental life.

ALLPORT, G. W., BRUNER, J. S., and JANDORF, E. M. "Personality under Social Catastrophe," pp. 347-66 in *Personality in Nature, Society, and Culture*, C. KLUCKHOHN and H. A. MURRAY (eds.). New York: Alfred Knopf, 1948.

Personal documents reporting the experiences of ninety individuals during the Nazi Revolution have provided a means whereby certain prevailing theories of behavior can be tested.

ALLPORT, G. W., and POSTMAN, L. *The Psychology of Rumor*. New York: Henry Holt & Co., 1947.

An analysis of "rumor" from the standpoint of the basic pattern of distortion, why rumors circulate, wartime rumors, and rumor in society.

BELL, J. E. *Projective Techniques*. New York: Longmans, Green & Co., 1948.

See earlier reference in this bibliography under General References in Motivation Research, page 403.

Brown, J. F. "A Modification of the Rosenzweig Picture-Frustration Test to Study Hostile Interracial Attitudes," *Journal of Psychology*, Vol. 24 (July, 1947), pp. 247-72.

The picture-frustration test as reported in this article seems to bring results that are not so deep an analysis of personality as may be obtained from the Rorschach or TAT techniques.

CLARK, K. B., and CLARK, M. P. "Racial Identification and Preference in Negro Children," pp. 169-78 in *Readings in Social Psychology*, T. M. NEWCOMB and E. L. HARTLEY (eds.). New York: Henry Holt & Co., 1947.

The problem of racial identification is related to the problem of the genesis of racial attitudes in children. This was tested by determining preferences for either a white doll or a brown doll by both southern and northern negroes. Results indicated a variety of motives for acceptance or rejection of the differentiated dolls.

DEWEY, J. *Human Nature and Conduct*. New York: Henry Holt & Co., 1948.

The book sets forth the belief that an understanding of habit and of different types of habit is the key to social psychology while the operation of impulses and intelligence gives the key to individualized mental activity.

DUBIN, S. S. "Verbal Attitudes Scores from Responses Obtained in a Projective Technique," *Sociometry*, Vol. 3 (January, 1940), pp. 24-48.

Various projective techniques are described and an evaluation of them is made on the basis of results obtained in the verbal attitude scores.

FARIS, G. "The Concept of Social Attitudes," pp. 3-16 in *Social Attitudes*, K. YOUNG (ed.). New York: Henry Holt & Co., 1931.

The various factors that influence human nature are discussed as an outgrowth of the modifications coming from biological, psychological, and cultural experiences.

FRENCH, V. V. "The Structure of Sentiments," *Journal of Personality*, Vol. 15 (June, 1947), pp. 247-82; Vol. 16 (September, 1947), pp. 78-108.

Sentiments have long been bound up with the problem of value. The genesis and development of sentiments and their relationship to attitudes is excellently explained in this article. The difficulties in measuring values and attitudes are well handled.

HAMMOND, K. R. "Measuring Attitudes by Error-Choice: An Indirect Method," *Journal of Abnormal and Social Psychology*, Vol. 43 (January, 1948), pp. 38-48.

Reproduced in Part II, Article 15, pp. 170-182.

HANFMANN, E. "Social Structure of a Group of Kindergarten Children," *American Journal of Orthopsychiatry*, Vol. 5 (October, 1935), pp. 407-10.

Results of this study showed how patterns of dominance and destruction motivate children and how play influences these patterns.

HOROWITZ, E. L. "The Development of Attitude toward the Negro," *Archives of Psychology*, No. 194 (January, 1936).

The results show how attitudes toward Negroes begin very early in the life span, develop gradually, are not innate but are formed by social forces.

HOROWITZ, E. L., and HOROWITZ, R. E. "Development of Social Attitudes in Children," *Sociometry*, Vol. 1 (January, 1938), pp. 301-38.

Data were collected to test hypotheses about the source of race attitudes in children. This material, together with consideration of the major body of data on the development of the attitudes, is discussed with reference to the theory of personality development.

JONES, A. W. *Life, Liberty and Property*. Philadelphia: Lippincott, 1941.

A report on 1,705 interviews of persons living in and around Akron relating to the rubber industry. The appendix is of considerable interest to marketing research workers because of the material contained on the interviewing process and the sample selection.

KRECH, D., and CRUTCHFIELD, R. S. *Theory and Problems of Social Psychology*. New York: McGraw-Hill Book Co., Inc., 1948.

The section (pages 205-72) on indirect methods of measuring the various characteristics of beliefs and attitudes is of use to marketing research workers.

LIKERT, R. "A Technique for the Measurement of Attitudes," *Archives of Psychology*, Vol. 22 (June, 1932), No. 140.

This is the original publication presenting Likert's scaling techniques as a substitute for Thurstone's more complicated approach, with evidence that the two yield essentially similar results. A basic work in the field.

LOEBLOWITZ-LENNARD, H., and RIESSMAN, F., JR. "A Preliminary Report on a Social Perception Test: A New Approach to Attitude Research," *Social Forces*, Vol. 24 (May, 1946), pp. 423-27.

The Social Perception Test is a projective test dealing with the problem of attitudes which can serve as an anti attitude detector (such as anti-Semitism or anti-Negro) and according to the author can enable us to analyze the value of effective education in altering perception and behavior.

———. "A Proposed Projective Attitude Test," *Psychiatry*, Vol. 9 (February, 1946), pp. 67-68.

Existing tools in social psychology are felt to have been unsuccessful in determining the nature of latent attitudes. The projective test described purports to get at this problem.

MACFARLANE, J. W. "Problems of Validation Inherent in Projective Methods," *American Journal of Orthopsychiatry*, Vol. 12 (July, 1942), pp. 405-12.

An interesting, critical analysis of the difficult conceptual and methodological problems presented by motivation research.

MCNEMAR, Q. "Opinion-Attitude Methodology," *Psychological Bulletin*, Vol. 43 (July, 1946), pp. 289-374.

A social scientist, outside of psychology, presents an appraisal of the techniques and methodologies utilized in studies of opinions and attitudes. The author concludes by saying that he fails to see any reason for optimism concerning war-born advances in the attitude opinion field of research.

MORENO, J. L. *Psychodrama*, Vol. 1. New York: Beacon House, 1946.

A comprehensive treatment of the psychodrama is found in this volume and the two which follow it. While it treats the therapeutic process involved in the psychodrama, it reveals some possibilities for role playing that may be useful to marketing workers.

MORGAN, J. J. B., and MORTON, J. T. "The Distortion of Syllogistic Reasoning Produced by Personal Convictions," *Journal of Social Psychology*, Vol. 20 (August, 1944), pp. 39-59.

Results showed that a person is likely to accept a conclusion which expresses his convictions with little regard for the correctness or incorrectness of the inferences involved.

MURPHY, G., and LIKERT, R. *Public Opinion and the Individual*. New York: Harper & Brothers, 1938.

This book presents a technique for the measurement of attitudes, personal backgrounds in relation to attitudes, and attitudes of the subjects five years later.

MURRAY, H. A., and MORGAN, C. D. "A Clinical Study of Sentiments," *Genetic Psychology Monographs*, No. I (August, 1945), pp. 1-149; No. II (November, 1945), pp. 153-309.

A useful discussion of sentiments and cathexis as well as feelings and affects related to dynamic environmental factors.

MURRAY, H. A., et al. *Explorations in Personality*. New York: Oxford University Press, 1938.

A report on clinical research on personality which sought evidence for the validity of certain general institutions about human nature as it operates in everyday life.

PROSHANSKY, H. M. "A Projective Method for the Study of Attitudes," *Journal of Abnormal and Social Psychology*, Vol. 38 (July, 1943), pp. 393-95.

Results of these investigations indicate that those strongly inclined toward extreme attitudes reveal their social orientation through their manner of report upon pictures of social conflict (for example, strongly prolabor or antilabor).

ROBINSON, D., and ROHDE, S. "Two Experiments with an Anti-Semitism Poll," *Journal of Abnormal and Social Psychology*, Vol. 41 (April, 1946), pp. 136-45.

The results reported in this study seem to support the hypothesis that different types of interviewers influence the degree of anti-Semitism expressed.

ROSENZWEIG, S. "The Picture-Association Method and Its Application in a Study of Reactions to Frustration," *Journal of Personality*, Vol. 14 (September, 1945), pp. 3-23.

A description is given of the picture-association method which seems to stand midway in design between the word association and the thematic apperception techniques.

SEELEMAN, V. "The Influence of Attitude upon the Remembering of Pictorial Material," *Archives of Psychology*, No. 258 (September, 1940).

The results reported in this study of Negro and white attitudes showed a pre-disposition to remember things about both groups in terms of favorable and unfavorable feelings toward them.

SHERIF, M., and CANTRIL, H. *The Psychology of Ego-Involvements*. New York: John Wiley & Sons, 1947.

An excellent discussion of attitude studies and methodological considerations and some experimental evidence concerning the basis of attitudes which are related to ego involvement.

SHOR, J. "Report on a Verbal Projective Technique," *Journal of Clinical Psychology*, Vol. 2 (July, 1946), pp. 279-82.

One of the objectives of the SIC (self-idea-completion) test is to encourage specifically freedom and spontaneity of response. The SIC test is developed along the lines of the traditional word-association test in which stimuli are beginnings of grammatical sentences selected to sample projective responses relating to general and specific feelings and attitudes. Marketing research workers will find it useful because it can be adapted to fit the individual needs of the test situation which is not found in many of the better known projective techniques.

SMITH, G. H. "Beliefs in Statements Labeled Fact and Rumor," *Journal of Abnormal and Social Psychology*, Vol. 42 (January, 1947), pp. 80-90.

An analysis was made of news-type statements which were labeled as "rumor," "facts," and no label at all. Results indicated that the "fact" label served as a tenuous standard of judgment while the "rumor" label was regarded as no label in that it constituted an ambiguous statement.

———. "The Relation of 'Enlightenment' to Liberal-Conservative Opinions," *Journal of Social Psychology*, Vol. 28 (August, 1948), pp. 3-17.

A report on so-called liberal and conservative students and their relative standing in both schooling and knowledge.

SOLLENBERGER, R. T., and PULFORD, E. H. "Attitudes of White Children toward Negroes in Social Situations as Measured by a Thematic Apperception Technique," *Inventory of Research in Racial and Cultural Relations*, Vol. 1, Bulletin No. 2 (1948), pp. 61-62.

This study seeks to contribute to an understanding of the genesis of racial prejudice and to test the hypothesis that "racial prejudices are learned."

SPENCER, D. "The Frankness of Subjects on Personality Measures," *Journal of Educational Psychology*, Vol. 29 (January, 1938), pp. 26-35.

A study showing that responses taken with signatures would result in evasions, resentments, and falsification of items.

SYMONDS, P. M. "Projective Techniques," pp. 583-90 in *Encyclopedia of Psychology*, P. L. HARRIMAN (ed.). New York: Philosophical Library, 1946.

A definition and discussion of a variety of projective techniques are found in this useful volume.

WESCHLER, I. R. "A Follow-up Study on the Measurement of Attitudes toward Labor and Management by Means of the Error-Choice Method: II," *Journal of Social Psychology*, Vol. 32 (August, 1950), pp. 63-69.

An objective treatment of the usefulness and limitations of the "error-choice" method of measuring attitudes toward labor and management.

———. "An Investigation of Attitudes toward Labor and Management by Means of the Error-Choice Method," *Journal of Social Psychology*, Vol. 32 (August, 1950), pp. 51-62.

Under the guise of an information test on labor-management relations, an inventory was constructed in which certain items were intended to elicit constant errors in the direction of the subject's known bias toward labor or management. The results achieved represented an effort to verify certain hypotheses concerning the "error-choice" method of attitude measurement.

WRIGHT, B. A. "An Experimentally Created Conflict Expressed by Means of a Projective Technique," *Journal of Social Psychology*, Vol. 21 (May, 1945), pp. 229-45.

Essentially, these investigations consist of experimentally creating a moral conflict and then observing whether that conflict is given expression through a projective technique.

12. CAN THE CLINICAL TECHNIQUES BE VALIDATED? BY DARRELL B. LUCAS

ERON, L. D. "Use of the Rorschach Method in Medical Student Selection," *Journal of Medical Education*, Vol. 29 (May, 1954), pp. 35-39.

The results of these investigations show that the use of the Rorschach test in the selection of medical students or in the prediction of their success in medical school is of little value.

HOLTZMAN, W. H., and SELLS, S. B. "Prediction of Flying Success by Clinical Analysis of Test Protocols," *Journal of Abnormal and Social Psychology*, Vol. 49 (October, 1954), pp. 485-90.

See reference in Part I, Article 6, page 405.

KELLY, E. L., and FISKE, D. W. *The Prediction of Performance in Clinical Psychology*. Ann Arbor: University of Michigan Press, 1951.

In 1947 a reassessment program was instituted to determine the validity of projective techniques as a prediction of performance. The assessments provide a sober evaluation of the results that can be achieved by four different projective techniques.

13. PROJECTIVE TECHNIQUES FROM AN ANALYTICAL POINT OF VIEW BY ROBERT FERBER

HATT, P. K. "Stratification in the Mass Society," *American Sociological Review*, Vol. 15 (April, 1950), pp. 216-22.

Sources of difficulty in determining the factor or combination of factors which delineate the social class are excellently discussed.

MCGUIRE, C. "Social Stratification and Mobility Patterns," *American Sociological Review*, Vol. 15 (April, 1950), pp. 195-204.

Marketing and motivation research workers will find this useful in evaluating certain aspects of the general theory of social stratification with special reference to status and mobility patterns in American society.

SANFORD, F. H., and ROSENSTOCK, I. M. "Projective Techniques on the Doorstep," *Journal of Abnormal and Social Psychology*, Vol. 47 (January, 1952), pp. 3-16.

A methodological study on the feasibility of projective techniques in door-to-door interviewing, finding that such interviews can be administered by people with no special psychological training, similar results are obtained by different coders, fully representative samples can be obtained with these techniques, and "controlled projection" and psychologically meaningful results are obtainable in this way.

14. THE THEMATIC APPERCEPTION TEST BY GARDNER LINDZEY

ATKINSON, J. W., and MCCLELLAND, D. C. "The Projective Expression of Needs: II. The Effect of Different Intensities of the Hunger Drive on Thematic Apperception," *Journal of Experimental Psychology*, Vol. 38 (December, 1948), pp. 643-58.

See reference in Part II, Article 14, page 415, for McClelland and Atkinson, "The Projective Expression of Needs: I."

BARTLETT, F. C. *Remembering*. London: Cambridge University Press, 1950.

A detailed account of experiments conducted in England on perceiving, imagining, and remembering. From these experiments a theory of remembering is developed.

BELLAK, L. "The Concept of Projection: An Experimental Investigation and Study of the Concept," *Psychiatry*, Vol. 7 (November, 1944), pp. 353-70.

A distinction is made between TAT techniques of projection and those of displacement.

———. "Thematic Apperception: Failures and the Defenses," *Transactions of the New York Academy of Science*, Vol. 12 (February, 1950), pp. 122-26.

A very brief summary of some of the limitations of Thematic Apperception Tests as well as a defense of the criticisms.

BRUNER, J. S. "Perceptual Theory and the Rorschach Test," *Journal of Personality*, Vol. 17 (December, 1948), pp. 157-68.

Understanding is better than believing, but measuring safer than understanding. The author shows how measurement in the strict sense has not proved overly applicable to the study of personality.

———. "Personality Dynamics and the Process of Perceiving," pp. 121-47 in *Perception, an Approach to Personality*, R. R. BLAKE and G. V. RAMSEY (eds.). New York: Ronald Press Co., 1951.

Some of the investigations related to motivational or behavioral factors as determinants of perception are excellently discussed in Chapter 5 of this book.

BRUNER, J. S., and POSTMAN, L. "An Approach to Social Perception," pp. 71-118 in *Current Trends in Social Psychology*, W. Dennis (ed.). Pittsburgh: University of Pittsburgh Press, 1948.

The study of perception is the study of how organisms come to see the world about them. Since the individual brings to a perceptual situation his prevailing motives, needs, attitudes, and personality structure, this book provides insight into how people perceive things.

———. "Tension and Tension Release as Organizing Factors in Perception," *Journal of Personality*, Vol. 15 (June, 1947), pp. 300-308.

This article serves as a good background to understanding perception factors as influenced by tension and tension release orientations.

COFER, C. N., and FOLEY, J. P., JR. "Mediated Generalization and the Interpretation of Verbal Behavior: I. Prolegomena," *Psychological Review*, Vol. 49 (November, 1942), pp. 513-40.

This paper reports on an attempt (1) to extend the objective principles derived from conditioned response experimentation to certain aspects of linguistic behavior, and (2) to show how such an interpretation of language behavior throws light on other psychological phenomena in which language plays an important role.

COX, B., and SARGENT, H. "TAT Responses of Emotionally Disturbed and Emotionally Stable Children: Clinical Judgment versus Normative Data," *Journal of Projective Techniques*, Vol. 14 (March, 1950), pp. 61-74.

The experiments reported were undertaken for the purpose of determining what actual differences exist between stable and disturbed members of one group of children, homogeneous with respect to age and grade. The results seem to indicate that error may result from uncritical generalization from pathological material alone.

- DOUGLAS, A. G. "A Tachistoscopic Study of the Order of Emergence in the Process of Perception," *Psychological Monographs*, Vol. 61, No. 6, Whole No. 287 (1947).

A detailed study of the larger problem of order and regularity of all natural events as manifested in the limited area of perception. Results indicated that perception is a process in time and space whose course spans an orderly sequence of developments.

- EDWARDS, A. L. "Political Frames of Reference as a Factor Influencing Recognition," *Journal of Abnormal and Social Psychology*, Vol. 36 (January, 1941), pp. 34-50.

An analysis of Freudian theory which involves pleasantness and unpleasantness. The results seem to show that an experience is excluded from consciousness not because it is unpleasant but rather because it involves conflict.

- ERON, L. D. "A Normative Study of the Thematic Apperception Test," *Psychological Monographs*, Vol. 64, No. 9, Whole No. 315 (1950).

A discussion of the materials, administration, technique of analyzing stories, statistical methods, and levels of interpretation that are involved in Thematic Apperception Tests.

- ERON, L. D.; TERRY, D.; and CALLAHAN, R. "The Use of Rating Scales for Emotional Tone of TAT Stories," *Journal of Consulting Psychology*, Vol. 14 (December, 1950), pp. 473-78.

A good discussion of the meaning and characteristics of the stimulus cards in TAT.

- FARBER, L. H., and FISHER, C. "An Experimental Approach to Dream Psychology through the Use of Hypnosis," *Psychoanalytic Quarterly*, Vol. 12 (April, 1943), pp. 202-16.

An excellent account of some of the limitations and experimental problems encountered when employing hypnosis as an analytical device.

- FRANCK, K. "Preferences for Sex Symbols and Their Personality Correlates," *Genetic Psychology Monographs*, Vol. 33 (1946), pp. 73-123.

This article points out the sex role as a motivational device in the identification process and how it affects personality.

- FRANCK, K., and ROSEN, E. "A Projective Test of Masculinity-Femininity," *Journal of Consulting Psychology*, Vol. 13 (August, 1949), pp. 247-56.

An understanding of the tests of the cultural factors in the male-female role is important to motivation research workers. This article explains a technique that is useful in scaling masculinity and femininity roles and which represents a continuum from projection to constructive self-representation.

- FREUD, S. "The Interpretation of Dreams," pp. 181-549 in *The Basic Writings of Sigmund Freud*, A. A. BRILL (ed.). New York: Modern Library, 1938.

This section of the book covers historical aspects of the literature on dream-problems up to 1900, the method of dream interpretation, the dream as wish-fulfillment, the material and sources of dreams, the dream-work, and the psychology of dream processes.

GRANDINE, L., and HARLOW, H. F. "Generalization of the Characteristics of a Single Learned Stimulus by Monkeys," *Journal of Comparative and Physiological Psychology*, Vol. 41 (October, 1948), pp. 327-38.

An investigation of the primary stimulus generalization in monkeys which indicated that training to a single stimulus resulted in a measurable amount of generalization to the stimuli from the training stimulus. This will be useful to motivation research workers in measuring the amount of advertising stimulation which might be generalized.

HAGGARD, E. A. "Experimental Studies in Affective Processes: I. Some Effects of Cognitive Structure and Active Participation on Certain Autonomic Reactions During and Following Experimentally Induced Stress," *Journal of Experimental Psychology*, Vol. 33 (October, 1943), pp. 257-84.

The present experiment was an investigation of some of the factors involved in both the establishment and extinction of autonomic reactions to a situation involving stress induced by strong electric shock. Results showed that those individuals who know most about the conditions involved in the situation and who took an active role in facing this experience showed less disturbance.

HARRISON, R. "Studies in the Use and Validity of the Thematic Apperception Test with Mentally Disordered Patients: II. A Quantitative Validity Study," *Character and Personality*, Vol. 9 (December, 1940), pp. 122-33.

Results of a picture test on 40 psychiatric hospital patients.

———. "Studies in the Use and Validity of the Thematic Apperception Test with Mentally Disordered Patients: III. Validation by the Method of 'Blind Analysis,'" *Character and Personality*, Vol. 9 (December, 1940), pp. 134-38.

An explanation of the "blind test" in which someone other than the person who is to analyze the materials administers the test so that the analyst is compelled to carry out his interpretations with no possibility of reliance on behavioral and similar clues.

HARRISON, R., and ROTTER, J. B. "A Note on the Reliability of the Thematic Apperception Test," *Journal of Abnormal and Social Psychology*, Vol. 40 (January, 1945), pp. 97-99.

This is of historical interest to motivation research workers since it indicates the ways in which general validity was established for Thematic Apperception Tests.

HENRY, W. E. "The Thematic Apperception Technique in the Study of Culture-Personality Relations," *Genetic Psychology Monographs*, Vol. 35 (February, 1947), pp. 3-135.

A good discussion of the TAT technique in the study of a person's ways of organizing experience so that the personality can project itself upon a plastic field of picture protocols and thus reveal the way an individual sees life.

HULL, C. L. *Principles of Behavior: An Introduction to Behavior Theory*. New York: Appleton-Century, 1943.

The nature of scientific theory and the way in which an objective theory of behavior can be formulated is covered in this book.

———. "The Problem of Primary Stimulus Generalization," *Psychological Review*, Vol. 54 (May, 1947), pp. 120-34.

An excellent study and reappraisal of the experimental facts related to the genuineness of primary stimulus generalization. Of interest to motivation research workers is the evidence presented here to indicate that early in the conditioning process primary generalization is nearly complete within the stimulus field originally conditioned.

KATZ, I. "Emotional Expression in Failure: A New Hypothesis," *Journal of Abnormal and Social Psychology*, Vol. 45 (April, 1950), pp. 329-49.

From a motivational point of view, an analysis of behavior in frustration must be derived from a knowledge or assumptions about the nature of frustration particularly as it relates to failure. The results of this test indicate that there is a connection between failure frustrations and the results on projective tests.

KORNER, A. F. *Some Aspects of Hostility in Young Children*. New York: Grune & Stratton, Inc., 1949.

A report on the use of TAT projective techniques to explain types of hostility and aggression that exist in young children. The results indicate that there is a causal relationship between the hostility behavior of children and the hostility expressions of their parents.

———. "Theoretical Considerations Concerning the Scope and Limitations of Projective Techniques," *Journal of Abnormal and Social Psychology*, Vol. 45 (October, 1950), pp. 619-27.

Although directed toward psychological studies, this discussion of the concepts underlying projective techniques and their limitations is well worth reading by commercial researchers as well. The difficulty of predicting behavior on the basis of such tests is discussed at some length.

KROUT, J. "Symbol Elaboration Test: The Reliability and Validity of a New Projective Technique," *Psychological Monographs*, Vol. 64, No. 4, Whole No. 310 (1950).

This monograph reviews the history, development, and use of the symbol elaboration test as a new projective technique. It is not advocated as a replacement for other tests (TAT and Rorschach) but as a supplement to tests now being used.

LASHLEY, K. S., and WADE, M. "The Pavlovian Theory of Generalization," *Psychological Review*, Vol. 53 (March, 1946), pp. 72-87.

This is a critical analysis of the Pavlovian Theory of Generalization which shows that the postulates upon which the theory is based are contrary to fact.

LEVINE, J. M., and MURPHY, G. "The Learning and Forgetting of Controversial Material," *Journal of Abnormal and Social Psychology*, Vol. 38 (October, 1943), pp. 507-17.

Previous studies have indicated the phenomenon of selective perception and recall; an individual notes and remembers material which supports his social attitudes better than material which conflicts with these attitudes. This investigation develops curves for the forgetting and learning processes.

LINDZEY, G. "An Experimental Examination of the Scapegoat Theory of Prejudice," *Journal of Abnormal and Social Psychology*, Vol. 45 (April, 1950), pp. 296-309.

This study purports to demonstrate that minority-group prejudice can be elevated, at least temporarily, by frustration experiences.

MAYMAN, M., and KUTNER, B. "Reliability in Analyzing Thematic Apperception Test Stories," *Journal of Abnormal and Social Psychology*, Vol. 42 (July, 1947), pp. 365-68.

This study attempted to determine the extent to which analyses of TAT stories made independently by two judges may agree. Interjudge agreements were determined significant as to the character with which the subject identified himself, the kind of tension-situation facing the identification-figure, and

other related factors that indicated the emotional involvement in producing the stories.

MCCLELLAND, D. C. "Measuring Motivation in Phantasy: The Achievement Motive," pp. 191-205 in *Groups, Leadership and Men: Research in Human Relations*, H. GUETZKOW (ed.). Pittsburgh: Carnegie Press, 1951.

The Thematic Apperception Test technique was used to provide an independent measure of achievement motivation in the research reported in this article. The research points to the possibility of measuring any motive independently.

MCCLELLAND, D. C., and ATKINSON, J. W. "The Projective Expression of Needs: I. The Effect of Different Intensities of the Hunger Drive on Perception," *Journal of Psychology*, Vol. 25 (April, 1948), pp. 205-22.

A report on experimental work related to the relationship between motive and precept as revealed by projective techniques of testing personalities. The results indicated that increasing intensities of hunger increased the food-related responses on projective tests.

MCCLELLAND, D. C.; CLARK, R. A.; ROBY, T. B.; and ATKINSON, J. W. "The Projective Expression of Needs: IV. The Effect of the Need for Achievement on Thematic Apperception," *Journal of Experimental Psychology*, Vol. 39 (April, 1949), pp. 242-55.

The results of this experiment seem to show that achievement as a motivating goal will influence the responses to TAT cards.

MCGINNIES, E., and BOWLES, W. "Personal Values as Determinants of Perceptual Fixation," *Journal of Personality*, Vol. 18 (December, 1949), pp. 224-35.

This article is a demonstration of the role of personal values in perceptual fixation, or learning, and suggests that perceptual behavior is characterized by three basic processes: selection, accentuation, and fixation.

MCGRANAHAN, D. V. "A Critical and Experimental Study of Repression," *Journal of Abnormal and Social Psychology*, Vol. 35 (April, 1940), pp. 212-25.

The results of the investigations conducted by the author seem to indicate that the relationship of repression to anxiety feelings is not the source of the neurotic condition, as is commonly held, but rather a product of it.

MILLER, N. E. "Theory and Experiment Relating Psychoanalytic Displacement to Stimulus-Response Generalization," *Journal of Abnormal and Social Psychology*, Vol. 43 (April, 1948), pp. 155-78.

The experiments and hypotheses described in this article were devised to relate the concepts of displacement and generalization into a theory of learning.

MURRAY, H. A. *Thematic Apperception Test Manual*. Cambridge, Mass.: Harvard University Press, 1943.

A description of the ways in which various Thematic Apperception Tests are handled by one of the originators of the TAT protocols.

PIOTROWSKI, Z. A. "A New Evaluation of the Thematic Apperception Test," *Psychoanalytic Review*, Vol. 37 (April, 1950), pp. 101-27.

A comparison of the Rorschach and TAT techniques is discussed along with nine rules of interpretation for the TAT.

POSTMAN, L. "Toward a General Theory of Cognition," pp. 242-72 in *Social Psychology at the Crossroads*, J. H. ROHRER and M. SHERIF (eds.). New York: Harper & Brothers, 1951.

Chapter 10 of this book covers the major trends in cognitive theory, the characteristics of a general cognitive theory, and the experimental evidence that has been developed in support of a system of cognitive hypotheses.

POSTMAN, L., and MURPHY, G. "The Factor of Attitude in Associative Memory," *Journal of Experimental Psychology*, Vol. 33 (September, 1943) pp. 228-38.

The problem of this study was to ascertain to what extent a student's attitude toward war will influence his memory for such material. Results seemed to indicate that acceptable events are retained much better than unacceptable ones.

RAPAPORT, D. "The Clinical Application of the Thematic Apperception Test," *Menninger Clinic Bulletin*, Vol. 7 (May, 1943), pp. 106-13.

An excellent article highlighting problems involved in the administration and interpretation of the TAT protocols.

RAZRAN, G. "Stimulus Generalization of Conditioned Responses," *Psychological Bulletin*, Vol. 46 (September, 1949), pp. 337-65.

This study points to a very crude and qualitative conditioned reflex gradient and confirms the existence of a true conditioned reflex generalization which is attributed to the organism's categorizing or rating of related stimuli on some sort of a crude similarity-dissimilarity scale.

RIESS, B. F.; Schwartz, E. K.; and Cottingham, A. "An Experimental Critique of Assumptions Underlying the Negro Version of the TAT," *Journal of Abnormal and Social Psychology*, Vol. 45 (October, 1950), pp. 700-709.

A study conducted to ascertain the influence of Negro and white figures when they are used on TAT cards. Also the effect of Negro and white administrators on Negro and white students was studied. The results seemed to show some differences in the length of protocol between Negro and white students, but this was not due to the type of figures on the TAT cards nor to whether the administrators were white or Negro.

ROSENZWEIG, S. "Apperceptive Norms for the Thematic Apperception Test: I. The Problem of Norms in Projective Methods," *Journal of Personality*, Vol. 17 (June, 1949), pp. 475-82.

A summary of the thinking in the field of TAT techniques with special reference to the lack of norms to apply in its scoring and interpretation (1949).

ROSENZWEIG, S., and FLEMING, E. E. "Apperceptive Norms for the Thematic Apperception Test: II. An Empirical Investigation," *Journal of Personality*, Vol. 17 (June, 1949), pp. 483-503.

A statement of the adult apperceptive norms for selected pictures of the 1943 published TAT series.

ROTTER, J. B. "Thematic Apperception Tests: Suggestions for Administration and Interpretation," *Journal of Personality*, Vol. 15 (September, 1946), pp. 70-92.

A practical treatment of ways in which TAT techniques can be used, administered, and interpreted.

SANFORD, R. N. "The Effects of Abstinence from Food upon Imaginal Processes: A Preliminary Experiment," *Journal of Psychology*, Vol. 2 (1936), pp. 129-36; "The Effects of Abstinence from Food upon Imaginal Processes: A Further Experiment," *Journal of Psychology*, Vol. 3 (1937), pp. 145-59.

The conclusions reached in these investigations support the view that food responses, which are assumed to be objectifications of food images, depend upon the strength of need for food, as measured in terms of elapsed time since food was last taken, and that where significant variations in food response frequency are observed, the presence of an underlying need tension is to be suspected.

SHIPOLA, E. M. "The Influence of Color on Reactions to Ink Blots," *Journal of Personality*, Vol. 18 (March, 1950), pp. 358-82.

The results of this experiment showed that the presence of hue in the blot did affect the responses obtained in the administration of the Rorschach test.

SYMONDS, P. M. *Adolescent Fantasy: An Investigation of the Picture-Story Method of Personality Study*. New York: Columbia University Press, 1949.

This study points to ways in which the picture-story (TAT) method can be used to penetrate into motives and underlying personality trends.

THOMPSON, C. E. "The Thompson Modification of the Thematic Apperception Test," *Rorschach Research Exchange*, Vol. 13 (December, 1949), pp. 469-78.

In using the TAT test with minority groups, the author noticed a dearth of material in the protocols given to the separate pictures. By including members of the minority groups in TAT pictures, better results were obtained, hence, suggesting a limitation that could be present in most TAT applications that are oriented toward certain well-known groups.

TOMKINS, S. S. *The Thematic Apperception Test; the Theory and Technique of Interpretation*. New York: Grune & Stratton, Inc., 1947.

This workbook about TAT is intended for those completely unfamiliar with TAT while also appealing to psychologists who have not entered this field of inquiry. Excellent for those who want to become familiar with the possibilities and limitations of TAT.

WEISSKOPF, E. A. "A Transcendence Index as a Proposed Measure in the TAT," *Journal of Psychology*, Vol. 29 (April, 1950), pp. 379-90.

The construction as well as the interpretation of TAT is frequently based on guesswork rather than on well-corroborated facts. This article indicates what groundwork will have to be done to place TAT upon a more solid empirical basis.

WICKENS, D. D. "Stimulus Identity as Related to Response Specificity and Response Generalization," *Journal of Experimental Psychology*, Vol. 38 (August, 1948), pp. 389-94.

The results of the tests suggest some tendency for response generalization, or the tendency for an act rather than a movement to become attached to a stimulus, to become a function of stimulus generalization.

15. THE ERROR-CHOICE TECHNIQUE BY KENNETH R. HAMMOND

EDWARDS, A. L. "Political Frames of Reference as a Factor Influencing Recognition," *Journal of Abnormal and Social Psychology*, Vol. 36 (January, 1941), pp. 34-50.

See reference in Part II, Article 14, page 412.

———. "The Retention of Affective Experiences: A Criticism and Restatement of the Problem," *Psychological Review*, Vol. 49 (January, 1942), pp. 43-53.

A restatement of previous ideas on learning and retention in which it was shown that experiences which conflict with our existing frame of reference will probably not cause a serious reorganization of the frame itself.

LEVINE, J. M., and MURPHY, G. "The Learning and Forgetting of Controversial Material," *Journal of Abnormal and Social Psychology*, Vol. 38 (October, 1943), pp. 507-17.

See reference in Part II, Article 14, page 414.

MCNEMAR, Q. "Opinion-Attitude Methodology," *Psychological Bulletin*, Vol. 43 (July, 1946), pp. 289-374.

A complete critical review of the opinion-attitude methodologies as used in 1946.

MYRDAL, G. *An American Dilemma*. New York: Harper & Brothers, 1948.

An excellent source of factual material on the marketing problems of the Negro market and the motivational background of these people.

NEWCOMB, T. M. "The Influence of Attitude Climate upon Some Determinants of Information," *Journal of Abnormal and Social Psychology*, Vol. 41 (July, 1946), pp. 291-302.

This study showed that the process of acquiring and retaining information relevant to a controversial issue is dynamically related to the process by which an attitude toward it is acquired.

PROSHANSKY, H., and MURPHY, G. "The Effects of Reward and Punishment on Perception," *Journal of Psychology*, Vol. 13 (1942), pp. 295-305.

The hypothesis of this investigation was that we learn to perceive in much the same way that we learn to act. The experimental subjects showed significant shifts in estimates of lengths and weights in the direction of the precepts which had been rewarded.

SHERIF, M. *The Psychology of Social Norms*. New York: Harper & Brothers, 1936.

An excellent book on the frame of reference in psychological phenomena, social values, basic needs and social values, and the formation of social norms in a group situation.

———. "A Study of Some Social Factors in Perception," *Archives of Psychology*, No. 187 (July, 1935).

This is a good study in social psychology which notes some social factors participating in the production of differential responses to the same stimulus.

SHERIF, M., and CANTRIL, H. "The Psychology of 'Attitudes,' Part II," *Psychological Review*, Vol. 53 (January, 1946), pp. 1-24.

The major thesis of this report and that of the previous report on Part I is that in spite of enormous variation in the content of attitudes, psychological principles of attitude formation are essentially the same, irrespective of what the attitude is concerned with.

WALLEN, R. "Ego-Involvement as a Determinant of Selective Forgetting," *Journal of Abnormal and Social Psychology*, Vol. 37 (1942), pp. 20-39.

The major finding of this study is that, when bogus ratings are presented as genuine, recalls of these ratings tend to be altered in such a way as to make them more compatible with subjects' opinions of themselves.

WHITE, R. K. "Value Analysis," *Journal of Social Psychology*, Vol. 19 (May, 1944), pp. 351-58.

This article covers the characteristics of the selection of basic value-words for describing and measuring free verbal expression.

ZUBIN, J. "Note on a Graphic Method for Determining the Significance of the Difference between Group Frequencies," *Journal of Educational Psychology*, Vol. 27 (September, 1936), pp. 431-44.

This is a report on a mathematical method of analysis that is applicable to all multiple-choice items. The use of an equation for chi square in terms of two compared frequencies having an elliptical relationship to each other is the method discussed. Visual analysis is made possible by this method.

16. THE RORSCHACH TEST BY DANIEL R. MILLER

BECK, S. J. *Rorschach's Test. Vol. I: Basic Processes*. New York: Grune & Stratton, Inc., 1950.

A complete coverage of processes used in administering the Rorschach test is given as disproof of the contention of certain intuitionists concerning the essential incommunicability of their skills.

CARP, A. L., and SHAVZIN, A. R. "The Susceptibility to Falsification of the Rorschach Psychodiagnostic Technique," *Journal of Consulting Psychology*, Vol. 14 (June, 1950), pp. 230-33.

This test indicates that, contrary to the results of Fosberg who said that the Rorschach test withstood all manipulation by the subjects, there are subjects who can manipulate their responses so as to make good or bad impressions.

CURTIS, H. S., and WOLF, E. B. "The Influence of the Sex of the Examiner on the Production of Sex Responses on the Rorschach," *American Psychologist*, Vol. 6 (August, 1951), pp. 345-46 (abstract).

The results of this test showed that there is a significant difference between male and female examiners on the number of records with sex responses. It appears necessary for each examiner to study his stimulus value and the resulting affect he has on the individuals he examines.

BELLS, K., et al. *Intelligence and Cultural Differences*. Chicago: University of Chicago Press, 1951.

This report is the first part of an extended study of cultural learning as it bears upon the solution of problems in mental tests. These tests may at times suggest certain motivational factors that may be intellectual in origin by relating cultural background to I.Q.

GIBSON, J. J. "A Critical Review of the Concept of Set in Contemporary Experimental Psychology," *Psychological Bulletin*, Vol. 38 (November, 1941), pp. 781-817.

The concept of set or attitude has never been clearly defined. Results of this test indicate that no common meaning can be discerned, but, instead, a number of ambiguities and contradictions have become evident.

HARROWER-ERIKSON, M. R., and STEINER, M. E. *Large Scale Rorschach Techniques*. Springfield, Ill.: Charles C. Thomas, 1945.

A manual for the Group Rorschach Test and the Multiple Choice Test which contains a discussion of methods, and interpretations of the results of these tests.

HAUGHT, B. F. "Mental Growth of the Southwestern Indian," *Journal of Applied Psychology*, Vol. 18 (February, 1934), pp. 137-42.

The findings of this study indicate that the Indians make lower scores than whites.

HERTZ, M. R. *Frequency Tables to Be Used in Scoring Responses to the Rorschach Ink-Blot Test*. 3d ed. Cleveland: Western Reserve University, 1946.

A useful manual on methods of interpreting ink-blot tests.

HERTZMAN, M. "A Comparison of the Individual and Group Rorschach Tests," *Rorschach Research Exchange*, Vol. 6 (July, 1942), pp. 89-108.

The results reported in this article indicate rather small differences between individual and group tests when all factors are considered. The feasibility of the employment of the group test for screening purposes is indicated.

JACOBI, J. *The Psychology of C. G. Jung*. New Haven: Yale University Press, 1951.

A concise presentation of the psychological theories of Jung in a manner that avoids technical particulars. A number of diagrams are included which aid in understanding certain functional relations.

JOEL, W. "The Interpersonal Equation in Projective Methods," *Rorschach Research Exchange*, Vol. 13 (December, 1949), pp. 479-82.

This paper presents a brief schematic discussion of the subject-examiner relationship, some suggestions regarding research, and a reorientation in the use and interpretation of projective methods.

KATZ, D. "Do Interviewers Bias Poll Results?" *Public Opinion Quarterly*, Vol. 6 (Summer, 1942), pp. 248-68.

A comparison between findings of white-collar interviewers with those of working-class interviewers. Though both interviewing staffs worked under the same instructions, they did not find the same public sentiment on labor and war issues.

KELLY, E. L., and FISKE, D. W. *The Prediction of Performance in Clinical Psychology*. Ann Arbor: University of Michigan Press, 1951.

In 1947 a reassessment program was instituted to determine the validity of projective techniques as a prediction of performance. The assessments provide a sober evaluation of the results that can be achieved by four different projective techniques.

KIMBLE, G. A. "Social Influence on Rorschach Records," *Journal of Abnormal and Social Psychology*, Vol. 40 (January, 1945), pp. 89-93.

The conclusion reached through these investigations would indicate that the Rorschach test is a somewhat unstable instrument because the protocol is easily susceptible to environmental influence.

KLINBERG, O. *Negro Intelligence and Selective Migration*. New York: Columbia University Press, 1935.

A report on the analysis of intelligence scores of Negroes who served in the army. It showed that length of residence in a favorable environment plays an important part in the intellectual level of the Negro children. Also, it was pointed out that the problem of selective migration varies with different communities.

KLOPFER, B., and KELLEY, D. M. *The Rorschach Technique*. Yonkers-on-Hudson, N.Y.: World Book Co., 1942.

A complete history and elaboration of the Rorschach method. The scoring problems as well as the interpretation problems are explained in detail.

LAZARUS, R. S. "The Influence of Color on the Protocol of the Rorschach Test," *Journal of Abnormal and Social Psychology*, Vol. 44 (October, 1949), pp. 506-16.

The experiment studied the contribution of color in the standard Rorschach series which aimed at determining whether the reaction pattern referred to as "color shock" is dependent upon the presence of color in the series. The results indicated that the concept of "shock" as induced by the presence of color in the slides is not supported.

LORD, E. "Experimentally Induced Variations in Rorschach Performance," *Psychological Monographs*, Vol. 64, No. 10, Whole No. 316 (1950).

A monograph on whether the administrator of a Rorschach test has a significant influence on the stimulus during the administration of projective tests. The results indicated that most frequent variations in Rorschach performance were associated with examiner differences.

LUCHINS, A. S. "Situational and Attitudinal Influences on Rorschach Responses," *American Journal of Psychiatry*, Vol. 103 (May, 1947), pp. 780-84.

The results reported in this study seem to indicate that present data for

"normal" subjects differ somewhat from the data obtained from soldiers and "normals" found in civilian populations.

MACALPINE, I. "The Development of the Transference," *Psychoanalytic Quarterly*, Vol. 19 (October, 1950), pp. 501-39.

This article focuses attention upon the lack of research and reasons for it in the explanation of transference as an integral part of psychoanalysis.

MACFARLANE, J. W. "Problems of Validation Inherent in Projective Methods," *American Journal of Orthopsychiatry*, Vol. 12 (July, 1942), pp. 405-12.

An interesting, critical analysis of the difficult conceptual and methodological problems presented by motivation research.

MATARAZZO, J. D., and MENSCH, I. N. "Reaction Time Characteristics of the Rorschach Test," *Journal of Consulting Psychology*, Vol. 16 (April, 1952), pp. 132-39.

This investigation adds to the normative data available to motivation research workers.

OBERNDORF, C. P. "Unsatisfactory Results of Psychoanalytic Therapy," *Psychoanalytic Quarterly*, Vol. 19 (July, 1950), pp. 393-407.

A constructive criticism of psychoanalysis as a diagnostic device, with emphasis upon nonstatistical approaches to this challenging problem. The research worker will find this article useful in gaining some insight into psychoanalytic therapy as an instrument of motivation research.

RICE, S. A. "Contagious Bias in the Interview: A Methodological Note," *American Journal of Sociology*, Vol. 35 (November, 1929), pp. 420-23.

An early and interesting report on how the bias on "drinking" in the mind of the interviewer was communicated by some process of suggestion to the mind of the interviewed as shown by responses of the latter.

ROBINSON, D., and ROHDE, S. "Two Experiments with an Anti-Semitism Poll," *Journal of Abnormal and Social Psychology*, Vol. 41 (April, 1946), pp. 136-45.

The results reported in this study seem to support the hypothesis that different types of interviewers influence the degree of anti-Semitism expressed.

ROGERS, C. R. "Where Are We Going in Psychology?" *Journal of Consulting Psychology*, Vol. 15 (June, 1951), pp. 171-77.

An excellent statement of the problems with which clinical psychology was faced in 1950.

SELLS, S. B. "Problems of Criteria and Validity in Diagnosis and Therapy," *Journal of Clinical Psychology*, Vol. 8 (January, 1952), pp. 23-28.

A critical appraisal is made of the limitations of clinical tests of personality. The author suggests that tests need to be standardized by reference to carefully selected populations representing specific life situations such as occupational, age, or other significant groups.

STANTON, F., and BAKER, K. H. "Interviewer Bias and the Recall of Incompletely Learned Materials," *Sociometry*, Vol. 5 (May, 1942), pp. 123-34.

Interviewers in the study were biased as to the responses expected by exposing them to a "key of correct" responses. Results showed that the bias exerts some determining effect on the outcome of the interview in spite of the fact that the interviewer is experienced.

WALLEN, R. "The Nature of Color Shock," *Journal of Abnormal and Social Psychology*, Vol. 43 (July, 1948), pp. 346-56.

The results of this test purported to show that the presence of color in the standardized Rorschach cards produced "shock" because it facilitates associa-

tions which have a disturbing effect upon the individual. (For contrasting views, see "The Influence of Color on the Protocol of the Rorschach Test" by R. S. Lazarus in Part II, Article 16.)

17. CARTOON AND PICTURE DEVICES BY MARTIN ZOBER

EHLE, E. L. "Techniques for Study of Leadership," *Public Opinion Quarterly*, Vol. 13 (Summer, 1949), pp. 235-40.

This article describes some of the pictorial and sentence-completion techniques which are being used in a study of leadership identification and acceptance.

MURRAY, H. A., *et al.* *Explorations in Personality*. New York: Oxford University Press, 1938.

See reference in Part II, Article 11, page 408.

ROSENZWEIG, S. "The Picture-Association Method and Its Application in a Study of Reactions to Frustration," *Journal of Personality*, Vol. 14 (September, 1945), pp. 3-23.

See reference in Part II, Article 11, page 408.

SMITH, G. H. *Motivation Research in Advertising and Marketing*. New York: McGraw-Hill Book Co., Inc., 1954.

A simple, well-written introduction to basic concepts in motivation research, with heavy emphasis on projective techniques and their application to commercial research.

18. THE MEANING OF GASOLINE SYMBOLS BY WILLIAM HENRY

No references listed.

19. PLAY TECHNIQUES FOR INTERVIEWING ON DURABLE GOODS BY C. W. GARBER, JR.

No references listed.

20. A SOCIOPSYCHOLOGICAL APPROACH TO CONSUMER BEHAVIOR BY WARREN J. BILKEY

BOARD OF GOVERNORS OF THE FEDERAL RESERVE SYSTEM. "Potential Buyers of Consumer Durable Goods in 1946," *Federal Reserve Bulletin*, Vol. 32 (August, 1946), pp. 847-48.

A detailed analysis of the characteristics of people who expected to buy automobiles and other consumer durables in 1946.

DAVIS, J. S. "Standards and Content of Living," *American Economic Review*, Vol. 35 (March, 1945), pp. 1-15.

An evaluation of the plane or content of living as a complex combination of consumption, working conditions, possessions, freedom, and "atmosphere," and the balance or harmony among them in relation to needs and wants. No data are presented in this article.

PART III. NONPROJECTIVE SURVEY TECHNIQUES

21. A QUALITATIVE APPROACH BY ALLAN R. WILSON

BRITT, S. H. "The Strategy of Consumer Motivation," *Journal of Marketing*, Vol. 14 (April, 1950), pp. 666-74.

A brief survey of the problems of classifying motives, of developing principles of motivation, of cutting across interdisciplinary lines, and of the need for more knowledge about people.

SKOTT, H. E. "Attitude Research in the Department of Agriculture," *Public Opinion Quarterly*, Vol. 7, No. 2 (Summer, 1943), pp. 280-92.

The importance in governmental routine of the execution of policies already legislated depends upon an understanding of how people are likely to react to specific measures and thus to improve the public relations aspects of both government agencies and consumers affected. Of interest to marketing research persons is the study given to the wishes and habits of farmers.

22. MEASURING CONSUMER ATTITUDES TOWARD PRODUCTS BY
GEORGE H. BROWN

No references listed.

23. WHY PEOPLE BUY PARTICULAR BRANDS BY SEYMOUR BANKS

BROWN, G. H. "Measuring Consumer Attitudes toward Products," *Journal of Marketing*, Vol. 14 (April, 1950), pp. 691-98.

A discussion of unique gains that can be achieved through the measurement of consumer attitudes toward products by means of unaided and aided recall tests of brand preferences for selected types of food products. The successive intervals (rank order) method showed that intensity of brand preferences can be measured and that the information provided can be used to focus marketing strategy and execution of product promotion in the right direction.

DURAND, D. *Risk Elements in Consumer Installment Financing: Technical Edition*. New York: National Bureau of Economic Research, 1941.

An excellent discussion of the application of discriminant functions to good and bad loan samples of consumers who use credit to purchase goods and services. The motivating factors studied in connection with consumers' marketing behavior as they relate to consumer loans are: qualitative attributes, like marital status and occupation, and quantitative attributes, like income and years of residence at the present address.

FISHER, R. A. "The Use of Multiple Measurements in Taxonomic Problems," *Annals of Eugenics*, Vol. 7 (September, 1936), pp. 179-88.

When two or more populations have been measured in several characteristics, special interest attaches to certain linear functions of the measurements by which the populations are best discriminated. This paper will be of interest to motivation research workers in discriminating between various factors that influence human behavior.

HOEL, P. G. *Introduction to Mathematical Statistics*. 2d ed. New York: John Wiley & Sons, 1954.

An excellent, simply written book especially recommended for the researcher desiring an acquaintance with mathematical statistics. It develops neatly and concisely frequency-distribution analysis and sampling and correlation theory. Can be read easily by anyone having a knowledge of elementary calculus.

JOHNSON, P. O. *Statistical Methods in Research*. New York: Prentice-Hall, Inc., 1949.

This book was devised primarily for graduate students in psychology and is particularly useful for discriminant analysis.

THURSTONE, L. L. "The Prediction of Choice," *Psychometrika*, Vol. 10 (1945), pp. 237-53.

This article is concerned with the concept of the discriminant dispersion in social measurement of such things as opinion polls, the measurement of attitudes, the measurement of political choices, and the measurement of consumer choices. For a good explanation of discriminant functions see H. Durant, "Experiences of Random (Probability) Sampling," *Public Opinion Quarterly*, Vol. 15 (Winter, 1951-52), pp. 765-66.

24. MEASURING THE TRUE STATE OF OPINION BY LOUIS GUTTMAN

GOODENOUGH, W. H. "A Technique for Scale Analysis," *Educational and Psychological Measurement*, Vol. 4 (Autumn, 1944), pp. 179-90.

A discussion of the tabulation technique, which enables scale analysis without a scalogram board and at the same time preserves its flexibility.

GUTTMAN, L. "A Basis for Scaling Qualitative Data," *American Sociological Review*, Vol. 9 (April, 1944), pp. 139-50.

An excellent treatment of scaling of populations, written in such a way that the layman can readily understand it. The uses and limitations of various methods are discussed.

GUTTMAN, L., and SUCHMAN, E. A. "Intensity and a Zero Point for Attitude Analysis," *American Sociological Review*, Vol. 12 (February, 1947), pp. 57-67.

This study gives a method of providing an invariant cutting point for an attitude that will be independent of the selection or "bias" of the specific question asked.

NOLAND, E. W. "Worker Attitudes and Industrial Absenteeism: A Statistical Appraisal," *American Sociological Review*, Vol. 10 (August, 1945), pp. 503-10.

This paper is concerned with a brief analysis of the nature of the industrial absentee problem, the rationale for the approach, and the statistical methodology that pointed toward the "causes" of absenteeism.

25. A STATISTICAL APPROACH TO MOTIVATION PROBLEMS BY J. B. LANSING

No references listed.

26. WHY TELEVISION COMMERCIALS SUCCEED BY HORACE S. SCHWERIN

No references listed.

PART IV. OTHER TECHNIQUES

27. WHY CONSUMER PURCHASES CHANGE OVER TIME BY RUTH P. MACK

No references listed.

28. WHY ADVERTISING READERSHIP VARIES BY DIK WARREN TWEDT

ADVERTISING RESEARCH FOUNDATION. *Continuing Study of Business Papers: No. 1, Automotive Industries*, issue of October 15, 1948. New York.

A useful readership survey of *Automotive Industries* magazine which employed a qualifying test and then the regular recognition method to measure the extent of readership of editorial and advertising items of the October 15, 1948 issue of the magazine.

———. *Continuing Study of Business Papers: No. 2, American Builder*, issue of February, 1950. New York.

A useful readership survey of *American Builder* magazine which employed a qualifying test and then the regular recognition method to measure the extent of readership of the February, 1950 issue of the magazine.

———. *Continuing Study of Business Papers: No. 3, American Machinist*, issue of March 6, 1950. New York.

A useful readership survey of *American Machinist* magazine which employed a qualifying test and then the regular recognition method to measure the extent of readership of editorial and advertising items of the March 6, 1950 issue of the magazine.

———. *Continuing Study of Business Papers: No. 4, Chemical Engineering*, issue of March, 1950. New York.

A useful readership survey of *Chemical Engineering* magazine which employed a qualifying test and then the regular recognition method to measure the extent of readership of editorial and advertising items of the March, 1950 issue of the magazine.

———. *Continuing Study of Farm Publications: No. 3, Successful Farming*, issue of May, 1947. New York.

A useful readership survey of *Successful Farming* magazine which employed a qualifying test and then the regular recognition method to measure the extent of readership of editorial and advertising items of the May, 1947 issue of the magazine.

FERGUSON, L. W. "The Importance of the Mechanical Features of an Advertisement," *Journal of Applied Psychology*, Vol. 19 (October, 1935), pp. 521-26.

An early study of the influence of such factors as size, page position, and day of publication upon the attention values of advertisements. Results seemed to show that these factors are opposite in their effects when compared with each other in metropolitan dailies and rural dailies. For example, metropolitan studies showed that these factors are very important in attention value, while in rural dailies there was no relationship.

FLESCH, R. "Measuring the Level of Abstraction," *Journal of Applied Psychology*, Vol. 34 (December, 1950), pp. 384-90.

Flesch discusses the problems of communication that are related to levels of abstraction, which can be estimated by computing a ratio of certain parts of speech to certain other parts of speech in written material. This ratio can be used as a measure of readability, either by itself or in combination with other elements.

LUCAS, D. B. "A Rigid Technique for Measuring the Impression Values of Specific Magazine Advertisements," *Journal of Applied Psychology*, Vol. 24 (December, 1940), pp. 778-90.

The method described is one of using published advertisements and unpublished advertisements in a test of reader recognition of advertisements. The recognition scores are adjusted to compensate for those who said they had seen unpublished advertisements.

LUCAS, D. B., and BRITT, S. H. *Advertising Psychology and Research*. New York: McGraw-Hill Book Co., Inc., 1950.

Parts V and VI take up various aspects of measuring the effect of advertising and gauging the size of audiences.

THORNDIKE, R. L. *Personnel Selection*. New York: John Wiley & Sons, 1949.

Problems of classification and selection as they relate to personnel work are excellently discussed. The statistical treatment is minimized to focus attention upon personnel problems.

THURSTONE, L. L. *Multiple-Factor Analysis*. Chicago: University of Chicago Press, 1947.

The purpose of this volume is to make available the results of psychometric work that has been done from 1937 to 1947 in the Psychometric Laboratory at the University of Chicago.

TWEDT, D. W. "A Table for Use with Flesch's Level of Abstraction Readability Formula," *Journal of Applied Psychology*, Vol. 35 (June, 1951), pp. 157-59.

The author suggests a method of taking Flesch's readability ratios and combining them with word length (syllables per 100 words) in a multiple re-

gression formula; the resulting index of readability correlates much higher with comprehension than does just the simple ratio suggested by Flesch. See also R. F. Flesch, *How to Test Readability*. New York: Harper & Brothers, 1951.

WOOLF, J. D. "It Isn't Size that Puts Pull in Advertising," *Advertising Age*, Vol. 22 (April 30, 1951), p. 43.

A series of reminiscences about the role of copy in producing advertising results. No quantitative information is offered to support the claims made in this article.

29. WHAT MAKES A BEST SELLER? BY JOHN HARVEY

BENNETT, E. A. *Fame and Fiction: an Enquiry into Certain Popularities*. London: Grant Richards, 1901.

A number of authors were called upon and a number of their works were analyzed to get a subjective appraisal of "the average reader and the recipe for popularity" around 1900.

BERREMAN, J. V. M. *Factors Affecting the Sale of Modern Books of Fiction: a Study of Social Psychology*. Unpublished Ph.D. dissertation, Department of Economics, Stanford University, 1940.

In contrast to many of the very early accounts of the way in which best sellers capture a market, this is a thoroughly objective account of the way in which best sellers become popular.

DURAND, D. *Risk Elements in Consumer Installment Financing: Technical Edition*. New York: National Bureau of Economic Research, 1941.

See reference in Part III, Article 23, page 423.

FISHER, R. A. "The Statistical Utilization of Multiple Measurements," *Annals of Eugenics*, Vol. 8 (August, 1938), pp. 376-86.

The author points out three independent lines of research on the treatment of multiple measurement. The methods are extended to the examination of collinearity and coplanarity of samples and to testing the significance of deviations.

———. "The Use of Multiple Measurements in Taxonomic Problems," *Annals of Eugenics*, Vol. 7 (September, 1936), pp. 179-88.

See reference in Part III, Article 23, page 423.

HICKS, G. "The Mystery of the Best Seller," *English Journal*, Vol. 23 (October, 1934), pp. 621-29.

A subjective approach to why *Anthony Adverse* was such a successful novel. This article deals with some of the psychological motives that may have been very influential factors in the success of this best seller.

LENROW, E. *Readers Guide to Prose Fiction*, pp. vii-xi. New York: Appleton-Century, 1940.

This reference is to the table of contents of this volume which provides a very interesting index to the various factors that influence human behavior and the ways in which they have found their way into literature. The same factors will also be found as important influences over marketing behavior.

KAPPEL, J. W. "Book Clubs and the Evaluation of Books," *Public Opinion Quarterly*, Vol. 12, No. 2 (Summer, 1948), pp. 243-52.

An attempt is made to evaluate some of the effects of the recent mushrooming of book clubs on the publishing of trade books as a whole. It was found that clubs do not lower literary standards, and they may raise them in some cases. To marketing research persons, the techniques used in making the evaluation are of special interest because a comparison was made of book club selections with best sellers.

KRIEG, L. L. "A Suggested Method of Analysing Children's Fiction Reading." Unpublished M.A. thesis, Graduate Library School, University of Chicago, 1943.

A limited study of the ways in which children project themselves into the books they read and thus identify themselves with various characters in them. Group patterns were found related to sex and maturity and may be of interest to motivation research workers from the point of view of the classification system suggested in this study.

SMITH, H. F. "A Discriminant Function for Plant Selection," *Annals of Eugenics*, Vol. 7 (November, 1936), pp. 240-50.

The nature of a plant may be expressed as a linear function of its characteristics by means of Fisher's concept of "discriminant functions." Hence, this is a treatment of the application of this method to plant selection and might be used to develop linear functions for motivational characteristics.

STEVENS, G. *Lincoln's Doctor's Dog, and Other Famous Best Sellers*. Philadelphia: Lippincott, 1938.

This book points out "how" best sellers are marketed and not "why" they are best sellers. A tremendously interesting account of how certain psychological appeals are exploited to make a "best seller" successful.

TIPPETT, L. H. C. *Random Sampling Numbers*. London: Cambridge University Press, University of London Tracts for Computers No. 15, 1927.

Contains 40,000 random sampling numbers.

WALLACE, N., and TRAVERS, R. M. W. "A Psychometric Sociological Study of a Group of Specialty Salesmen," *Annals of Eugenics*, Vol. 8 (May, 1938), pp. 266-302.

The broad purpose of this study, of which the first section is briefly reported in this article, was preliminary investigation into the development of a technique for team study by anthropologists, economists, geneticists, medical workers, psychologists, sociologists, and statisticians, of individuals in groupings under present-day conditions. Advanced techniques today for a study of this nature make it apparent that this is of primarily historical interest to motivation research workers.

WEEKS, E. "What Makes a Book a Best Seller?" *New York Times Book Review*, Vol. 41 (December 20, 1936), pp. 2, 15.

A subjective appraisal of one editor's viewpoint, giving reasons why he feels that *Gone with the Wind* was a successful novel.

30. WHY DAILY DEPARTMENT STORE SALES FLUCTUATE BY A. T. STEELE

CONTROLLERS CONGRESS. 1947 *Departmental Merchandising and Operating Results of Department Stores and Specialty Stores*. National Retail Dry Goods Association, 1948.

A quantitative analysis of the data that marketing research workers will find useful in working out studies involving department and specialty stores.

CROXTON, F. E., and COWDEN, D. J. *Practical Business Statistics*. 2d ed. New York: Prentice-Hall, Inc., 1948.

A simplified version for business students of the text on *Applied General Statistics* by the same authors.

DEAN, J. "Department-Store Cost Functions," *Studies in Mathematical Economics and Econometrics*, O. LANGE et al. (eds.) Chicago: University of Chicago Press, 1942, pp. 228 ff.

An excellent analysis of cost functions of department store operations with special emphasis given to problems of collecting, rectifying, and analyzing the data in order to find the relation between cost and output, with other influences being held constant.

MCNAIR, M. P.; GRAGG, C. I.; and TEELE, S. F. *Problems in Retailing*. New York: McGraw-Hill Book Co., Inc., 1937.

The chapter on personnel and organization indicates how adequate selling personnel in a department store can be very helpful in motivating consumers to buy.

VAN CLEEF, E. "The Influence of Weather on Street Car Traffic in Duluth, Minnesota," *Geographical Review*, Vol. 3 (February, 1927), pp. 126-34.

A prophetic article on the degree to which weather influences not only street car traffic but also the business traffic. Many of the factors that are discussed in this article, such as temperature, amount of rain and snow, and barometric pressure, have become commonplace items in determining the forecasts of trade volumes of many types.

"Weather Profit," *Business Week* (March 30, 1946), pp. 72 ff.

A popular treatment of the ways in which weather conditions (wind, rain, humidity, temperature, for example) exert a definite influence upon housewives who go shopping. This idea is of considerable use to motivation research workers in the prediction of human behavior in the market place.

31. WHY A MAIL-ORDER CATALOG IS SUCCESSFUL BY GEORGE KINGSLEY ZIPF

BRYANT, W. B.; WASHBURN, S. L.; and OUTERBRIDGE, D. G. "Some Psychological Determinants of the Structure of Advertising in a Classified Telephone Directory," *American Journal of Psychology*, Vol. 61 (October, 1948), pp. 540-44.

The authors use the classified section of the December, 1947, Boston Classified Telephone Directory to study the relationship of number of ads and their sizes. They found that there is an inverse relationship between number and size of ads.

CONVERSE, P. D. *Retail Trade Areas in Illinois*. Urbana, Ill.: University of Illinois Bulletin, Vol. 43, Business Studies No. 4, 1946.

A study that includes the factors which influence consumers' buying habits and buying motives in various trade areas in Illinois.

———. *A Study of Retail Trade Areas in East Central Illinois*. Urbana, Ill.: University of Illinois Bulletin, Vol. 41, Business Studies No. 2, 1943.

A quantitative method of ascertaining the breaking point of trade between cities. This takes into account the use of geographical factors that motivate shoppers to trade in a given urban area.

REILLY, W. J. *Methods for the Study of Retail Relationships*. Austin: University of Texas Bulletin, No. 2944, 1929.

An explanation of the various ways in which the "Law of Retail Gravitation" can be applied to newspaper circulation, charge accounts, and distribution of types of stores by size of towns.

STEWART, J. Q. "Demographic Gravitation: Evidence and Applications," *Sociometry*, Vol. 11 (1948), pp. 31-57.

A mathematical treatment and refinement of the work done by W. J. Reilly in his study of retail relationships. The author suggests in this article that the relations of people to one another require more than people to support them: they require also natural resources and technological facilities. The proportions of

these in terms of numbers of people involved, resources, and technological facilities is expressed in a formula of demographic gravitation.

———. "Empirical Mathematical Rules Concerning the Distribution and Equilibrium of Population," *The Geographical Review*, Vol. 37 (July, 1947), pp. 461-85.

Marketing research workers will find the discussion of mathematical methods applied to social phenomena of real interest. The author designates this as "social physics" and advocates the adaptation of social studies to natural principles.

STROHKARCK, F., and PHELPS, K. "The Mechanics of Constructing a Market Area Map," *Journal of Marketing*, Vol. 12 (April, 1948), pp. 493-96.

A study sponsored by the Curtis Publishing Company that investigated the Reilly and Converse formulas for the breaking point of trade between cities. This is a concrete example of quantification of information related to shopping lines of goods and the market area these goods command around trading centers.

WELLINGTON, A. M. *The Economic Theory of the Location of Railways*, 6th ed. New York: John Wiley & Sons, 1906.

Tables are presented which show the operating results that must be considered in the location of railroads in order to forecast successful and profitable operation.

ZIPF, G. K. "Brand Names and Related Social Phenomena," *American Journal of Psychology*, Vol. 63 (July, 1950), pp. 342-66.

An interesting study of the use of brand names and trade names in the distribution of goods and the relationship between the words used in trade names and the same underlying principles that seem to govern the selection and usage of words in an everyday cultural vocabulary. The results indicated that the relationships in the use of everyday words and those found for brand names were the same.

———. *Human Behavior and the Principle of Least Effort*. Cambridge, Mass.: Addison-Wesley Press, 1949.

From a marketing and motivation research point of view, this text reveals how the principle of the even distribution of minimized work over time might be utilized in the solution of a variety of distribution problems.

———. "A Note on Brand-Names and Related Economic Phenomena," *Econometrica*, Vol. 18 (July, 1950), pp. 260-63.

The author reports a number of instances with respect to brand names which fit his earlier hypothesis: $y^2 = c/x$.

INDEX

A

A & P, 32
 Adamic, L., 296 ff.
 Adams, A. J., 190
 Adkins, M. M., 167
 Adler, A., 13
Advertising Age, 76, 359
 Advertising Research Foundation, 65, 122, 124, 130-31, 348 ff.
 Aided recall, 270 ff.
 Airplane travel, reasons for, 10, 139
 Alderson, Wroe, 3, 11 ff.
 Allies, attitude toward, 121
 Allport, G. W., 115, 120
 Altman, C. H., 194
American Builder, 349 ff.
American Council on Education Studies, 189
 American Dairy Association, 130
 American Documentation Institute, 351
American Economic Review, 239
 American Institute of Public Opinion, 187
American Journal of Orthopsychiatry, 111, 112, 149, 154, 189, 196
American Journal of Psychiatry, 184
American Journal of Psychology, 399, 402
American Journal of Sociology, 187
 American Marketing Association, 122, 251, 315
 American Metereological Society, 388
American Psychologist, 188, 190
American Sociological Review, 293, 315
 Anderson, G. L., 62
 Anderson, H. H., 62, 112
 Animal experiments, 153
Annals of Eugenics, 378
Archives of Psychology, 104, 107, 164, 171
 Argument completion, 118
 Arluck, E., 185
 Arnold, M. B., 145
 Atkinson, J. W., 148, 160-61, 162
 Audit Bureau of Circulation, 349
 Auld, F., 127

B

Baker, K. H., 187-88
 Balken, E. R., 145, 161
 Ballantine beer, 327
 Banks, S., 247, 275, 277 ff.
 Barnard, M. M., 378

Bartlett, F. C., 164
 Baruch, D. W., 110
 Bayer aspirin, 33
 Bayes Theorem, 136
 Beck, S. J., 185
 Beers, N. R., 383
 Behaviorism, 12-13, 19
 Belden, J., 33
 Belden, Joe, and Associates, 33
 Bell, J. E., 110, 113, 116
 Bellak, J., 151, 160, 169
 Bennett, E. A., 367
 Benton, A. L., 185
 Bernberg, R. E., 75, 103 ff., 247
 Berreman, J. V. M., 363, 364, 371, 380
 Berry, F. A., Jr., 383
 Bias, in collecting data, 278-79
 Bibring-Lehner, G., 187
 Bilkey, W. J., 233 ff.
 Blake, R. R., 144, 159
 Blanchard, M. B., 194
 Blatz beer, 327
 Blueberry Fill Pie Mix, 101
 Bogardus, E. S., 104
 Bolgar, H., 112
 Bollay, E., 383
 Book of the Month Club, 363, 372
 Borden Co., 323
 Borden's milk, 98
 Bottome, P., 370
 Boulding, K., 20
 Bourland, D. D., Jr., 401
 Bowles, W., 164
 Brand recognition, 27
 Brill, A. A., 163
 Brisco, N. B., 386
British Journal of Medical Psychology, 185
 Britt, Steuart, 75, 89 ff., 258, 359-60
 Brown, G. H., 247, 268 ff., 277, 286
 Brown, J. F., 109-10
 Brown, J. S., 153
 Brown, L. O., 75, 76 ff.
 Bruner, J. S., 120, 144, 148, 149, 150, 159
 Bryant, W. B., 399
 Budweiser beer, 327
 Buehler, C., 112
 Buhler, C., 186
 Buhler, K., 186
 Buick automobile, 7, 8, 37, 45
 Bureau of Labor Statistics, 241-42
 Burney, R. C., 158-59

Business Week, 358, 360, 388

Buying habits, 12-13, 338 ff.

C

Cadillac automobile, 10, 45, 125

Calden, G., 192-93, 195

Callahan, R., 157

Campbell, P., 239

Cannell, C., 3, 4 ff.

Cantril, H., 171, 172

Carmichael, L., 164

Carp, A. L., 184

Cartoon technique, 109, 197 ff.

Cayton, H. R., 309

Chain interview, 35

Character and Personality, 161

Chase and Sanborn instant coffee, 325

Chein, I., 111, 148

Chi square test, 370

Chicago Tribune, 36, 206

Chicago, University of, 270

Chrysler automobile, 45

Cities Service gasoline, 208 ff.

Clark, K. B., 111, 164

Clark, M. P., 111

Clark, R. A., 160-61, 162

Cleveland, S. E., 192

"Client-Centered Therapy," 19

Coca-Cola, 51-52

Cofer, C. N., 153, 189

Cohen, L., 192-93, 195

Color shock, 183

Combs, A. W., 20

Comic Strip Sequence, 34

Communism, attitude toward, 117

Congressional Record, 401

Conoco gasoline, 208 ff.

Conrad, H. S., 194

Continuous panel, 234 ff., 268 ff.

Continuous Studies of Readership, Advertising Research Foundation, 348 ff.

Controllers Congress, 386

Converse, P. D., 391

"Cornell Technique" for scalogram analysis, 294 ff.

Cost and Profit Outlook, Alderson and Sessions, 11

Cottingham, A., 165

Cowden, D. J., 383

Cox, B., 161

Cox, R., 388

Croxton, F. E., 383

Crutchfield, R. S., 105, 121

Curtis, H. S., 188

D

Dancer-Fitzgerald-Sample, 76

Davis, J. S., 239

Dean, J., 383

Del Monte peaches, 96, 97, 99

Department of Agriculture, U.S. Govt., 259

Depth interviews, 9, 10, 27-28, 36, 58-61, 66-68, 77-79, 126, 133, 136-37, 202, 205, 255

Dewey, J., 104

Dichter, E., 3, 21 ff., 31, 50, 69

Direct questioning, 30, 54, 94, 105, 123, 130, 134

Diven, K., 153

Doll play, 111, 112

Douglas, A. G., 144

Drake, St. C., 309

Draw-a-Man test, 194

Dreams, 155

Dubin, S. S., 112

DuBois, C., 270, 284

Durand, D., 378

D-X gasoline, 208 ff.

E

Econometrica, 402

Educational and Psychological Measurement, 293

Educational Psychology Monographs, 183

Edwards, A. L., 164, 172

Eells, K., 193

Eiconics, 20

Eisenhower, D. D., 124

Electric shock, 153-54

Elizur, A., 192

English Journal, 363

Eron, L. D., 127, 157, 158

Error-choice technique, 119, 120, 170 ff.

Esty, W., 124

Evans, M. C., 111

Experience Appraisal Blank, 114

Experimental design, 139

F

Farber, L. H., 155

Faris, G., 104

Federal Reserve Board, 318

Federal Reserve Board Index of Department Store Sales, 387

Federal Reserve Bulletin, 243

Ferber, R., 75, 133 ff.

Ferguson, L. W., 360

Fiction Catalog, 372

Field workers, 123, 137

Filer, R. N., 193

Fischer, L., 112

Fisher, C., 155

Fisher, R. A., 378

Fiske, D. W., 127, 189

Fleming, E. E., 158, 165

Flesch, R., 351

Fold-over Technique, 305-6, 308-9

Foley, J. P., Jr., 153

Ford automobile, 7, 8, 37, 49, 319
 Forgetting, selective, 172
 Four-picture Test, 144
 Franck, K., 155
 Frank, L. K., 189
 Free association tests, 10, 209-10
 French, V. V., 107
 Frenkel-Brunswik, E., 107
 Freud, S., 13, 21, 58-60, 63, 154, 155, 163, 184, 187
 Freudian, 18, 21, 63, 68, 156, 232
 Frigidaire refrigerator, 57
 Fromm, E., 13, 17
 Fromme, A., 109
 Frustration, 188
 Frustration, social, 151-52

G

Gallup polls, 348
 Garber, C. W., Jr., 231 ff.
 Gasoline symbols, meaning of, 206 ff.
 General Electric refrigerator, 57
 General Mills, 270, 323
Genetic Psychology Monographs, 107, 112, 149, 155
Geographical Review, 383
 George Washington instant coffee, 325
 Gerould, K. F., 367
 Gerson, M. J., 186, 189
 Gestalt psychology, 13-15
 Gibby, R., 196
 Gibson, J. J., 185
 Glover, E., 60
Gone with the Wind, 363
 Grandine, L., 153
 Grassi, J. R., 186, 189
 Grogg, C. I., 386
 Group interviewing, focussed, 67, 79-81
 Guetzow, H., 162
 Guided association, 210
 Guilford, J. P., 127, 188
 Gulf gasoline, 62
 Guttman, L., 247, 248, 293 ff.

H

Hackett, A., 370, 372
 Haggard, E. A., 153
 Haire, M., 35, 75, 93 ff.
 Haldeman-Julius, E., 362
 Hammond, K. R., 75, 119, 170 ff.
 Hanfmann, E., 112
 Harlow, H. F., 153
 Harriman, P. L., 110
 Harrison, R., 152, 161, 162
 Harrower-Erikson, M. R., 186
 Hartley, E. L., 111
 Hartmann, G. W., 172
 Hartshorne, H., 104
Harvard Business Review, 248

Harvey, J., 337, 361 ff.
 Haught, B. F., 193
 Hazel Bishop lipstick, 330
 Henry, W., 206 ff.
 Henry, W. E., 149, 162, 165
 Herodotus, 23
 Hertz, M. R., 185
 Hertzman, M., 185
 Hicks, G., 363, 366-67
 Hogan, H. P., 164
 Holtzman, W. H., 61, 128
 Homburger, E., 111
 Hopi Indians, 162, 165
 Horney, K., 13, 17
 Horowitz, E. L., 111, 112
 Horowitz, R. E., 106, 112
 Houghton, 33
 Hull, C. L., 153
 Hunger, effect of, 147-48
 Hutt, M. L., 186, 196
 Hypnosis, 155, 162, 186, 189, 190

I

Identification, 152
 Institute for Motivational Research, Inc., 21
 Institute of Social Research, 317
 "Instrumental view," 15-17
 Intensity analysis, 304 ff., 312-15
International Journal of Opinion and Attitude Research, 103
International Journal of Psycho-Analysis, 187
 Interpretation, 69-70, 87-88, 91, 93, 137-38, 145 ff., 260-61
 Interviewer bias, 139, 187, 233
 Interviewer fatigue, 271
Inventory of Research in Racial and Cultural Relations, 109, 111

J

Jacobi, J., 190
 Jandorf, E. M., 120
 Jew-Gentile relations, 107, 108, 110, 114-15, 187
 Joel, W., 185
 Johnson, H., 383
 Johnson, S., 359
 Jones, A. W., 118
 Jones, H. E., 194
Journal of Abnormal and Social Psychology, 108, 114, 119, 128, 144, 145, 151, 152, 153, 164, 165, 166, 170, 172, 182, 183, 185, 187
Journal of Applied Psychology, 117, 172, 193, 348, 349, 351
Journal of Clinical Psychology, 117, 186
Journal of Comparative and Physiological Psychology, 153

- Journal of Comparative Psychology*, 153
Journal of Consulting Psychology, 127, 155, 157, 183, 184, 190
Journal of Educational Psychology, 114, 177, 194
Journal of Experimental Psychology, 153, 160, 161, 164
Journal of General Psychology, 189, 190
Journal of Marketing, 35, 89, 93, 197, 258, 268, 277, 381, 391
Journal of Medical Education, 127
Journal of Mental Science, 112
Journal of Personality, 107, 109, 144, 149, 154, 157, 158, 163, 164, 197
Journal of Projective Techniques, 161, 196
Journal of Psychology, 106, 107, 109, 145, 148, 157, 161, 162, 172
Journal of Social Psychology, 109, 112, 115, 116, 119, 120, 148, 172, 177, 233
Jung, C. G., 13, 17, 190

K

- Kamman, G. R., 196
Kappel, J. W., 363
Karchin, S. J., 190
Kardiner, 13
Katz, D., 187
Katz, I., 148
Kelley, D. M., 185, 186
Kelly, E. L., 127, 189
Kelly, G., 112
Kelvinator refrigerator, 57
Kenmore range, 399
Kimble, G. A., 185
Klebanoff, S. G., 161
Klein, D. B., 155
Klein, L. R., 318
Klineberg, O., 193, 194
Klopper, B., 185, 186
Kluckholm, C., 120
Koffka, K., 13
Kohler, W., 13
Korner, A. F., 166, 168
Kreck, D., 105, 121
Krieg, L. L., 367
Krout, J., 155
Kutner, B., 152
Kuznets, G., 177

L

- Labor-Management attitudes, 119-20, 173 ff.
Lacey, J. I., 127
Lange, O., 383
Langley, A. L., 374-76
Lansing, J. B., 315 ff.
Lashley, K. S., 153
Lazarus, R. S., 154, 183

- Lefever, O. W., 186
Lenrow, E., 366
Levine, J. M., 164, 172
Levine, K. N., 186, 189
Levine, R., 148
Levy, D. M., 154
Lewin, K., 233-34, 243
Lieberman, A. M., 163
Lie detector test, 134
Likert, R., 104, 113
Lincoln automobile, 45
Lindzey, G., 143 ff., 151, 161
Linear discriminant function, 287 ff.
Linear regression function, 287 ff.
Link, H. C., 33
Literary Guild, 363, 372
Loebowitz-Lennard, H., 108, 116
Lord, E., 188
Lowenfeld, M., 112
Lucas, D. B., 75, 122 ff., 349, 359-60
Luchinus, A., 184

M

- Macalpine, I., 185
MacCorquadale, K., 145
Mack, R. P., 338 ff.
"Madison Ave.," 21
Mail questionnaires, 253
Make-A-Picture-Story test, 144
Margulies, H., 185
Martineau, P., 3, 36 ff.
Masserman, J. H., 145, 161
Matarazzo, J. D., 183
Maxwell House coffee, 96, 97, 98, 99, 100, 325
May, M. A., 104
Mayman, M., 152
McCleary, R. A., 154
McClelland, D. C., 148, 158-59, 160-61, 162
McFarlane, J. W., 62, 110, 189
McGinnies, E., 164
McGranahan, D. V., 164
McIntyre, F., 383
McNair, M. P., 386
McNemar, Q., 103-4, 106, 170
Meehl, P. E., 72, 145
Meltzoff, J., 190
Meninger Clinic Bulletin, 156
Mensch, I. N., 183
Miles Laboratories, 323
Military, attitude toward, 117
Miller beer, 327
Miller, D. R., 182 ff.
Miller, N. E., 153
Miller, R. B., 167
Milton, E. O., 196
Minnesota Multiphasic Personality Inventory, 194

- Mobilgas gasoline, 208 ff.
Monographs of the Society for Research in Child Development, 167
 Moreno, J. L., 113-14
 Morgan, C. D., 107, 108, 117, 118
 Morgan, J. J. B., 115
 Morton, J. T., 115
 Motivation, psychology of, 8-9
 Motivation Committee, Advertising Research Foundation, 122, 130-31
 Mott, F. L., 362
 "Movie Story Game," 111
 Muller, H., 367
 Multiple factor analysis, 350 ff.
 Multiple regression equations, 357-58, 382 ff.
 Multiple validation, 26-27
 Murphy, G., 113, 148, 164, 172
 Murray, H. A., 107, 108, 117, 118, 120, 148, 149, 150, 162, 163, 166, 168, 198
- N**
- Nash automobile, 135
Nation, 372, 374
 National Analysts, 130
 National Research Council, 233
National Research Council Bulletin, 233
 National Socialism, attitude toward, 120, 121
 Navaho Indians, 162, 165
 Negro-white differences, 106-7, 108, 109, 110, 111, 165
 Nescafé instant coffee, 35, 96, 97, 98, 99, 100, 101, 102, 325
New York Herald Tribune, 364, 372, 374
New York Post, 69
New York Times, 363, 372, 374
New Yorker, 372, 374
 Newcomb, T. M., 111, 233
 Nielsen, A. C., 323

O

- Oberndorf, C. P., 187
 Observational Studies, 27
 Open questions, 28
 Orwell, G., 367
 Outerbridge, D. G., 399

P

- Pabst beer, 327
 Paired comparisons, 271-72
 Parents, attitude toward, 107, 117
 Parliament cigarettes, 39
 Pavlovian, 153
 Pearson, K., 370
 Pearsonian coefficient, 385
 Personification, 34-35
 Phelps, K., 391

- Philco gasoline, 208 ff.
 Pictorial tests, 106-7, 109, 130, 133, 148, 155, 197 ff.
 Picture recognition, 67
 Picture response, frustration, 85-86, 197 ff.
 Picture selection test, 107-8
 Pillsbury flour, 98
 Piotrowski, Z. A., 149, 150, 166
 Play techniques for interviewing, 231 ff.
 Play therapy, 111
 Plymouth automobile, 39
 Politz, Alfred, Research, 3, 50 ff.
 Pontiac automobile, 39
 Postman, L., 115, 148, 149, 150, 159, 164
 Pottharst, K., 196
Printers' Ink, 31, 89
 Projective techniques, 10, 27-28, 61-64, 75, 81-86, 91-93, 93 ff., 103 ff., 122, 126, 129, 130-31, 133 ff., 143 ff., 170, 183, 233, 321
 Proshansky, H. M., 108, 172
Psychiatry, 116, 151, 184, 190
 Psychiatry, clinical, 77, 113, 127-28
 Psychoanalysis, 13-15, 17, 29, 58-60, 123, 126, 144
Psychoanalytic Quarterly, 155, 185, 187
Psychoanalytic Review, 149
 Psychoanalytic therapy, 187
 Psychodrama; *see* Role-playing
Psychological Bulletin, 103, 104, 112, 143, 145, 153, 161, 185
Psychological Monographs, 144, 155, 157, 188
Psychological Review, 145, 153, 171, 172
 Psychology, clinical, 5-6, 56, 61-62, 66-68, 95, 111-12, 122 ff., 131-32, 143 ff., 178, 182 ff., 320
 Psychology, experimental, 185
 Psychology, laboratory, 5, 171, 172
 Psychology, social, 104-5, 171, 172
 Psychometrics, 92
Psychosomatic Medicine Monographs, 162
Public Opinion Quarterly, 50, 103, 187, 231, 259, 361, 363
Publishers Weekly, 370, 372, 373, 374
 Pulford, E. H., 108-9
 Pumpian-Mindlin, E., 29
 Pure Oil gasoline, 208 ff.

Q

- Qualitative market research, 248 ff.
 Questions, unstructured, 116, 126, 133

R

- Rabelais, F., 21
 Rafferty, J. E., 157
 Ramsey, G. V., 159
 Random sample, 252-53, 254
 Random selection, 389-90

- Rank, O., 13, 17
 Rank ordering, 271
 Rapaport, D., 156
 Rating methods, 247
 Rating scale, numerical, 278 ff.
 Razran, G., 153
 RCA television, 319
 RCA Victor Corp., 323
Reader's Digest, 362
Readers Guide, 372, 374
 Recall, 107, 116, 119, 172
 Regression analysis, 315 ff.
 Reilly, W. J., 391
 Reliability, 90-91, 93, 136, 176, 180-81, 232, 327-30, 368-70, 380
 Reliability coefficients, 180
 Religion, attitude toward, 107, 117
 Remington electric shaver, 325
 Renaud, H., 161
 Revere pots, 342
 Rice, S. A., 91, 187
 Riess, B. F., 165
 Riessman, F., Jr., 108, 116
 Robinson, D., 114, 187
 Robinson, O. P., 386
 Roby, T. B., 158-59, 160-61, 162
 Rodnick, E. H., 161
 Rogers, C. R., 19, 20, 190
 Rohde, S., 114, 187
 Rohrer, J. H., 159
 Role-playing, 34, 38, 106, 110-11, 113
 Rorschach, H., 182
 Rorschach Institute, 127, 134
 Rorschach Research Exchange, 144, 165, 185, 186, 192, 196
 Rorschach test, 34, 67, 95, 106, 127, 128, 133, 144, 182 ff., 319
 Rosen, E., 155
 Rosenzweig, S., 109, 158, 165, 197
 ROTC, 161
 Rotter, J. B., 152, 157
 Rumford's baking powder, 96, 97
 Rumors, belief in, 116
 Rumors, transmission of, 115-16
 Russia, attitude toward, 116, 173 ff.
- S
- Sampling, 21 ff., 55, 71, 87, 123-24, 126, 131, 134, 140-41, 142, 167, 199, 208-9, 258, 278, 294-95, 308, 368-69, 370-71
 Sanders, R., 192
 Sanford, R. N., 107, 148, 160, 167, 168
 Sarbin, T. R., 189
 Sargent, H., 161
Saturday Review of Literature, 372, 374
 Scaling methods, 247
 "Scalogram" analysis, 293 ff.
 Schachtel, E. G., 184
 Schachtitz, E., 157
 Scheier, F. T., 64
 Schick electric shaver, 325
 Schlitz beer, 327, 328
 Schultz, H., 243
 Schwartz, E. K., 165
 Schweppes Tonic Water, 39
 Schwerin, H. S., 321
 Scriven, L. E., 64 ff.
 Sears, R. R., 149
 Sears-Roebuck, 388 ff.
 Securities and Exchange Act of 1933, 266
 Seeleman, V., 107
 Sells, S. B., 61, 128, 186
 Semantics, 88
 Sentence completion, 33-34, 38, 54, 67, 84-85, 117, 130-31, 133, 134, 157
 Sentiments examination, 117
 Sequence analysis, 145
 Sex, attitude toward, 107, 117, 168
 Sexes, difference between, 155, 165, 188
 Shakow, D., 194
 Shavzin, A. R., 184
 Shell gasoline, 62, 208 ff.
 Sherif, M., 159, 171, 172
 Shor, J., 117, 196
 Shuttlesworth, F. K., 104
 Significance, tests of, 379-80
 Siipola, E. M., 144
 Sinclair gasoline, 208 ff.
 Singer, J. L., 190
 Skott, H. E., 259
 Smirnoff vodka, 39
 Smith, G. H., 116, 198
 Smith, H. F., 378
 Smith, M., 344
 Snygg, D., 20
Social Forces, 108
 Society for Projective Techniques, 124
 Sociometry, 113
Sociometry, 112, 391
 Sociopsychology, 6-8
 Socio-psychological approach, 233 ff.
 Sollenberger, R. T., 108-9
 Spearman-Brown formula, 177, 179, 181
 Spencer, D., 114
Springfield Republican, 372, 374
 Standard gasoline, 208 ff.
Standard Rate and Data Service, 351
 Stanford-Binet test, 194
 Stanton, F., 187-88
 Starch surveys, 348
 Statistical guides, external, 252
 Steele, A. T., 337, 381 ff.
 Stein, M. I., 144, 162
 Steiner, M. E., 186
 Stevens, G., 362
 Stewart, J. Q., 391
 Stone, R. G., 383

Story Association test, 211
 Story plots, invention of, 118, 144
 Stouffer, S. A., 293
 Strohkarch, F., 391
 Structural validation, 27
Successful Farming, 358, 360
 Successive intervals method, 271
 Suchman, E. A., 315
 Sullivan, H. S., 13, 190, 191
 Sunbeam electric shaver, 325
Survey of Consumer Finance, 1953, 318
 Survey Research Center, 134, 318
 Swift's, 85
 "Symbolic view," 17-18
 Symonds, P. M., 104, 110, 167, 168

T

Tabulation, 199
 Teele, S. F., 386
 Tendler, A. D., 117
 Terry, D., 157
 Texaco gasoline, 208 ff.
 Thematic Apperception Test, 34, 38, 95-96, 107, 108-9, 116, 118, 128, 143 ff., 194, 197 ff.
 Thom, H. C. S., 386
Thomas' Register, 390, 402
Thomas' Register of American Manufacturers, 401
 Thompson, C. E., 165
 Thompson, J. Walter, Advertising Agency, 359
 Thorndike, E. L., 172
 Thurston's centroid method, 355
Time Magazine, 65, 372, 374
 Time series analysis, 134
 Tippet, L. H. C., 368
 Tomkins, S. S., 154-55, 169
 Toni home permanent, 323, 328
Transactions of the New York Academy of Science, 169
 Transcendence index, 157
 Travers, R. M. W., 183, 378
 Tri-Dimensional Apperception test, 144
 Tryon, R. C., 177
 Tubdenham, R. D., 62
 Twedt, D. W., 348 ff.
Twentieth Century Authors, 372

U

U.S. Army Air Forces, 127, 188
 U.S. Army Service Forces, Research Branch, 293, 294, 308
 U.S. Dept. of Commerce, 344
 U.S. Strategic Bombing Survey, 121

V

Validation, 26-27, 60-63, 91-92, 93, 122, 125-26, 128, 130-32, 138, 139-40, 143, 147, 155, 177, 185, 232
 Van Cleef, E., 383
 Verbal tests, 106, 114 ff.
 Veterans Administration, 127
 Vicary, J. M., 3, 31 ff., 69-70

W

Wade, M., 153
 Wallace, N., 378
 Wallen, R., 172, 183
 Walter, A. A., 164
 War, attitude toward, 107, 109, 117
 War Bonds, 9
 Warren, 374-76
 Washburn, S. L., 399
 Watson, J. B., 12
 Watson, W. S., 172
 Watsonian, 18
 Weeks, E., 363, 367
 Weisskopf, E. A., 157
 Wellington, A. M., 390
 Wertheimer, M., 13
 Weschler, I. R., 75, 103 ff., 119, 120, 247
 White, R. K., 177
 White, R. W., 162
 Wickens, D. D., 153
 Williams, R. J., 61
 Williams, W. L., 190
 Wilson, A. R., 248 ff.
 Wilson, G. P., Jr., 144
 Winston, F., 344
 Wolf, E. B., 188
 Wonder bread, 96, 97
 Wood, A., 185
 Wood, A. J., and Co., 3, 64
 Woolf, J. D., 359
 Word association methods, 27, 32-33, 67, 82-84, 116-17, 129-30
 "World test," 112
 Wright, B. A., 112
 Written tests, 106, 114 ff.
 Wulfeck, W. H., 124, 131
 Wyatt, F., 145

Y

Yntema, T. O., 383
 Younker Bros., Inc., 382 ff.

Z

Zeitschrift der Psychologie, 172
 Zillig, M., 172
 Zipf, G. K., 337, 388 ff.
 Zober, M., 197 ff.
 Zubin, J., 61, 177

This book has been set on the Linotype in 12 and 10 point Garamond #3, leaded 1 point. Selection titles are in 12 point Garamond #3 Roman caps. The size of the type page is 27 by 46½ picas.



